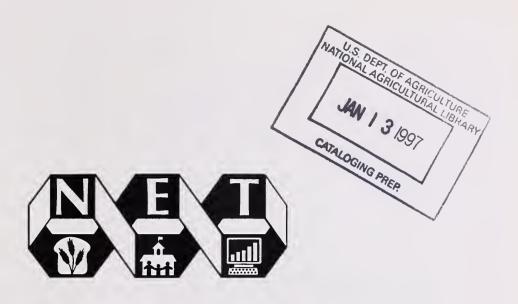
Historic, Archive Document

Do not assume content reflects current scientific knowledge, policies, or practices.





Evaluation Guide for the Nutrition Education and Training Program

United States Department of Agriculture Food and Consumer Service

April 1995



Credits

This Guide was developed by USDA under contract with the National Food Service Management Institute (Dr. Josephine Martin, Executive Director)

Lumina Training Associates Dr. Ann Robinson, R.D., Director Dr. Gail Trapnell, Associate

The following persons served as technical reviewers of the Guide:

Staff of Nutrition and Technical Services Division, Food and Consumer Service, United States Department of Agriculture

Patricia Daniels Jane Mandell Dr. Martha Poolton

Staff of Regional Offices, FCS, USDA

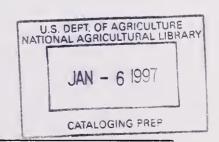
Michele Bouchard, Mid-Atlantic Regional Office Dr. Robert Dean, Midwest Regional Office Stella Nash, Mountain Plains Regional Office Sandra Remancus, Northeast Regional Office Nena Bratianu, Southeast Regional Office Joi Hatch, Southeast Regional Office Louise Lapeze, Southwest Regional Office Mavis Buchholz, Western Regional Office

Executive Board, National Association of State NET Coordinators (NASNET)

Roberta Richey, President Meg Wagner, Vice President Jacqueline Frederick, Secretary/Treasurer Gordon Odermann, Legislative Coordinator Blanche Harrison, Immediate Past-President

Regional Liaisons, NASNET

Linda Van Rooy, Maryland, Mid-Atlantic Region
Barbara Kalina, Minnesota, Midwest Region
Jeanne Puerta, Wyoming, Mountain Plains Region
Susan Fiore, Connecticut, Northeast Region
Helen Minns Derryberry, Tennessee, Southeast Region
Deborah Simpson, Texas, Southwest Region
Carolyn Garland, Nevada, Western Region



An Equal Opportunity Program - NET is available to all persons without regard to race, color, national origin, age, sex, or disability. Persons who believe they have been denied equal opportunity for participation should write to the Secretary of Agriculture, Washington, D.C. 20250.

i

Table of Contents



	Preface		V
•	Chapter I:	Introduction to Evaluation	1
	-	Purpose	
		Overview of the Guide	
		What is Evaluation?	
		Role of the NET Coordinator	
		Why Evaluate?	
		The Evaluation Challenge	
		Summary	
Δ'Δ	Chapter II:	Developing and Managing Evaluation	5
		Evaluation as Part of the A-P-B-I-E Process	
		Planning for Evaluation	
		A balance between the use of formative and summative evaluation	
		A balance between formal and informal evaluation design	
		A balance between quantitative and qualitative evaluation methods	
		Fair evaluation in relation to proximity of outcomes	
		Managing Evaluation	
		Step 1. Develop an evaluation plan for each objective	
		in the State Plan	17
		△ Evaluation questions for each objective	
		For each evaluation question, identify data to be collected	
		For each evaluation question, identify	
		a standard to assess acceptability	14
		For each evaluation question, identify an	
		appropriate data collection method or technique	15
		△ A work schedule with assigned responsibilities	
		Step 2. Determine how the evaluation plan will be managed	
		Step 3. Monitor the evaluation plan	
		Step 4. Analyze and interpret results of the evaluation	
		Step 5. Use the evaluation findings in management decision making	
		Summary	
		·	
\bigoplus	Chapter III:	The CIPPI Model for Evaluation	22
	-	Characteristics of an Evaluation Model	
		The CIPPI Model	2
		Summary	2



₩	Chapter IV:	Application of the CIPPI Model - NET Evaluation Plan	26
	Chapter V:	Explanation of Context Evaluation	33
	onup to:	Introduction	
		Why Address this Type of Evaluation?	
		When to Use Context Evaluation	
		How to Design Context Evaluation	
		Summary	
•••	Chapter VI:	Explanation of Input Evaluation	37
		Introduction	37
		Why Address this Type of Evaluation?	37
		When to Use Input Evaluation	37
		How to Design Input Evaluation	38
		Summary	39
***	Chapter VII	: Explanation of Process Evaluation	40
	r	Introduction	
		Why Address this Type of Evaluation?	
		When to Use Process Evaluation	
		How to Design Process Evaluation	
		Summary	
	Chapter VII	I: Explanation of Product Evaluation	45
	•	Introduction	
		Why Address this Type of Evaluation?	
		When to Use Product Evaluation	
		How to Design Product Evaluation	
		Summary	
¥	Chapter IX:	Explanation of Impact Evaluation	50
	<u>.</u>	Introduction	
		Why Address this Type of Evaluation?	
		When to Use Impact Evaluation	
		How to Design Impact Evaluation	
		Summary	



<u></u>	References	.53
	Appendices	.59
	Appendix A: Guidelines for Developing Evaluation Instruments	.60
	Appendix B: Data Collection Method: Case Study	.63
	Appendix C: Data Collection Method: Participant	
	Reaction Questionnaire	.65
	Appendix D: Data Collection Method: Portfolio Assessment	.68
	Appendix E: Sampling Protocol for NET Needs	
	Assessment Surveys	.70
	Appendix F: Excerpts from the NET Strategic Plan	.77
≘ \	Glossary	.79
X	Index	.81



Preface

Evaluation is an integral part of the management of the Nutrition Education and Training Program (NET). The purpose of this Guide is to assist the State NET Coordinator in planning and managing the evaluation component of the State Plan as required in the NET regulations (7 CFR 227.37(b)(14)).

Evaluation is a creative process and there are many different approaches to evaluation design and methodology. The authors and reviewers of this Guide had to make decisions about what to include in the Guide and how to describe an approach to evaluation for use by State NET Coordinators. The approach to evaluation described here is used by some experts in the field, however, there are other valid approaches to evaluation and other useful methods and techniques. The NET Coordinators and other readers are encouraged to design evaluation using this Guide as well as other references and experts who are available for consultation on specific aspects of evaluation. To design effective and efficient evaluation, the NET Coordinator must be sensitive to the information needs and resource constraints of the NET Program and the individual projects for which evaluation is designed.

This Guide is written as a companion document to the Needs Assessment Guide For The Nutrition Education And Training Program (1994). Because needs assessment and evaluation are so closely related in program design and management, there are necessarily many areas of common content in the two Guides. To the extent possible, repetition of content in the NET Evaluation Guide has been eliminated by referencing appropriate parts of the NET Needs Assessment Guide. Therefore, it is recommended that the reader have a copy of the NET Needs Assessment Guide easily available for reference.

Note S

849





Chapter I Introduction to Evaluation

Purpose

The purpose of this Evaluation Guide is to provide guidance to State agencies for designing the evaluation component of the Nutrition Education and Training (NET) State Plan required by Food and Consumer Service (FCS) regulations and defined by the Strategic Plan for Nutrition Education (1993). The use of a comprehensive evaluation model as described in this Guide can enhance the effectiveness of a State's Child Nutrition Program.

Overview of the Guide

This Guide has been developed to assist the NET Coordinator in planning and managing the evaluation of the NET State Plan which details the individual projects and activities funded through NET. It is recommended that the Guide be used as a reference along with the NET Needs Assessment Guide (1994). This Guide begins with an explanation of evaluation and reasons for program evaluation in Chapter I, Introduction to Evaluation. Chapter II, Developing and Managing Evaluation, describes how to develop an evaluation plan for the NET State Plan. Chapter III, The CIPPI Model for Evaluation, presents one of the most commonly used frameworks for the design of a comprehensive evaluation plan. Subsequent chapters provide detailed information on how to use the Model for NET evaluation. Chapter IV, Application of the CIPPI Model - NET Evaluation Plan, provides an example of the evaluation component of some State Plan objectives. Chapter V, provides an Explanation of Context Evaluation. An Explanation of Input Evaluation is provided in Chapter VII is an Explanation of Process Evaluation. An Explanation in Chapter IX.

It is recommended that the reader review Chapter IV, Application of the CIPPI Model - NET Evaluation Plan before going further in this Guide. This example of an evaluation plan for some NET State Plan objectives illustrates the evaluation model described in this Guide and will make the content more meaningful to the reader.

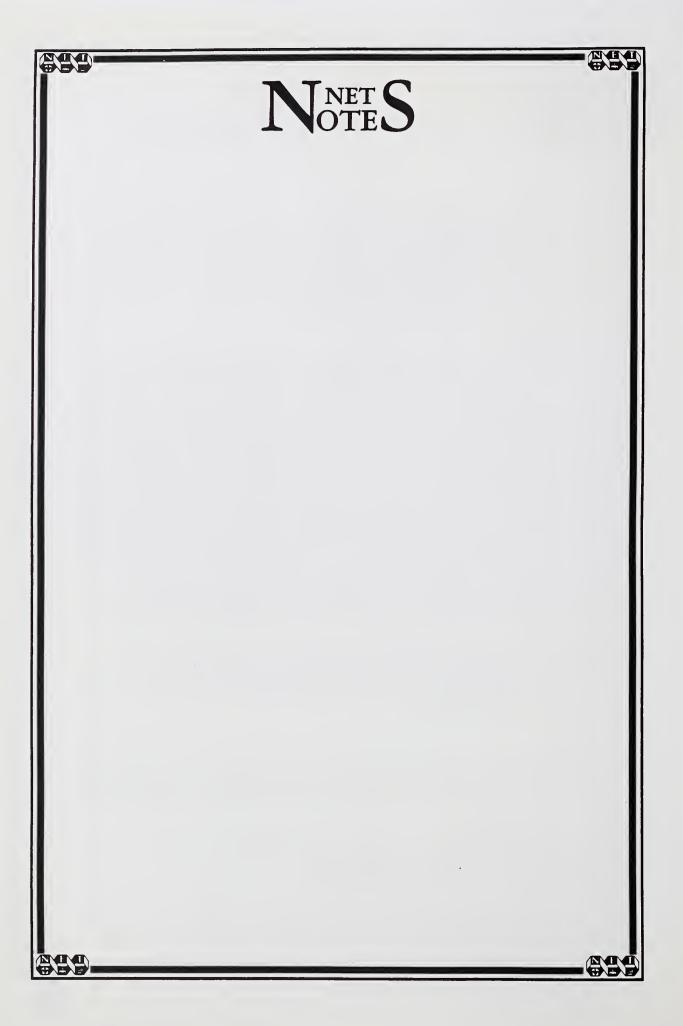
What is Evaluation?

Evaluation of the NET Program involves determining the worth or value of specific resources and activities in meeting the needs for which the Program is intended. Evaluation could be designed to assess:

- the State NET Program as a whole,
- S a food and nutrition training program component of the State Plan, or
- a nutrition education project component of the State Plan.

The worth or value of the Program is determined through quantitative and qualitative measurements of program services including the quality of the services provided and the success of the







services in achieving the desired outcomes. The results of evaluation are used for policy and management decision making and for program accountability (Stufflebeam, 1971; Wholey, 1991).

Program evaluation may be viewed as a prospective or a retrospective approach. A NET evaluation for decision making could be considered a prospective approach, in this case it is essentially a needs assessment process. The information gathered will relate to the future and be used to answer questions about what the NET Program should be doing and how the services should be delivered. For example, the evaluation of a statewide, elementary nutrition education program will provide valuable information about the design of future elementary nutrition education programs.

A NET evaluation designed for accountability could be considered a *retrospective* approach, or a look back at things past. Here, the evaluator is looking back on decisions that were made and the resources and activities offered to determine if the defined needs were met. An evaluation of a NET Program that addresses whether the State agency is making a difference over the long term, provides data for accountability.

Role of the NET Coordinator

As the manager of the State NET Program, the NET Coordinator is involved in every aspect of the Program from needs assessment to project management to program evaluation. Evaluating the program goals and the specific projects funded by the NET Program is one of the most challenging aspects of the NET Coordinator's role.

Why Evaluate?

To effectively manage a State NET Program, consideration should be given to the following reasons for program evaluation.

1. NET regulations require evaluation to be addressed in the State Plan.

FNS regulations 7 CFR Part 227 describe the implementation and operation requirements of the NET Program. The regulations (Section 227.37(b) (14)) specify that the NET State Plan shall include, "...plans to evaluate program activities including an evaluation component for each objective of the State Plan." Program evaluation must be addressed in order to effectively manage the Program and to use the available funds for the maximum impact in the State. The NET regulations describe the required evaluation:

227.31(c) Evaluations. The State agency shall conduct formal evaluations of program activities at least annually. These evaluations shall be aimed at assessing the effectiveness of the various activities undertaken by the State and local agencies. State officials shall analyze why some activities have proved effective while others have not and shall initiate appropriate improvements. Representatives of State Advisory Councils and agencies that participate in the State agencies' programs as well as representatives of other groups shall be invited to participate in the evaluations. The results of the evaluations shall be used to make adjustments in ongoing activities and to plan activities and programs for the next year's State plan. The State agency shall submit a plan for evaluation of Program activities as part of the State plan in accordance with 227.37(b)(14).

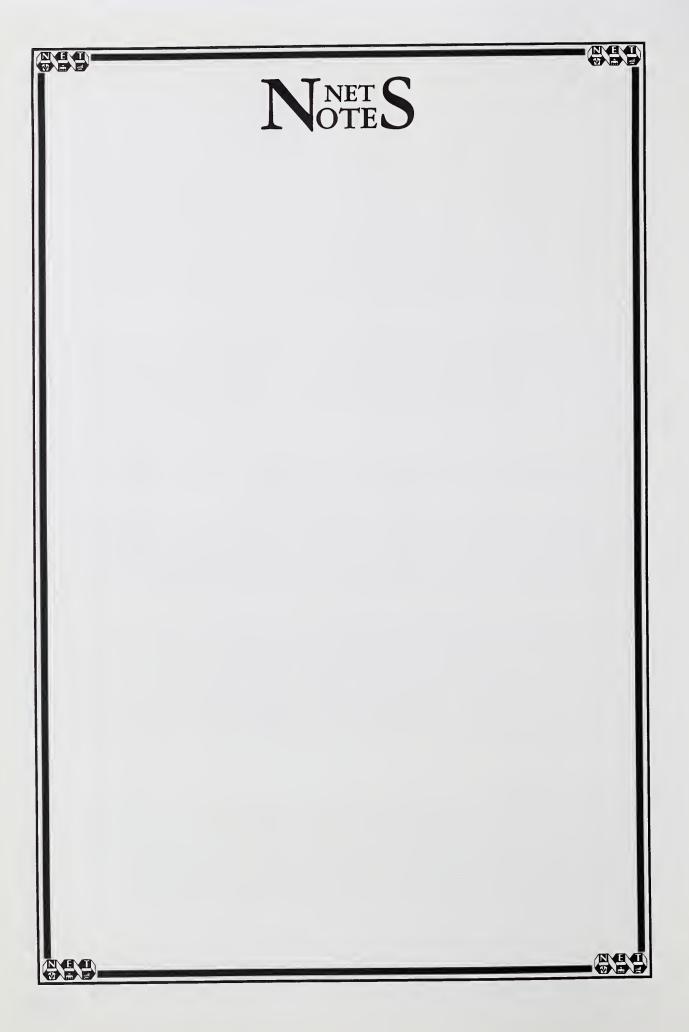




- 2. The Strategic Plan For Nutrition Education (1993) defines evaluation as a basis for program development, improvement, and accountability. The Nutrition and Technical Services, Food and Nutrition Service (now Food and Consumer Service), USDA, initiated a strategic planning process for the Nutrition Education and Training Program in March 1992. The participants in the strategic planning process defined a vision and strategic direction for the NET Program that provides a clear and concise conception of the program's future. The strategic direction helps guide program planners as they make decisions, "...regarding organizational policy, strategy, programs, and resource allocations." (Mandell, 1993). The Strategic Direction addresses a State's responsibility for conducting "...formal needs assessments and evaluations as the basis for program development, improvement, and accountability." The ten goals included in the Strategic Plan are national planning goals that should be used by States to develop their own goals as defined by the identified needs of the targeted populations (See Appendix F).
- 3. Evaluation data should be used to make decisions for program improvement. Assessment data provides the NET Coordinator with information on which to base program planning and design decisions, while evaluation data provides information for program improvement and accountability. Program goals and objectives provide the basis on which to measure program improvement.
- 4. Evaluation can help to show cost-effectiveness. Evaluation data can be used to help justify funding for the NET Program and other Child Nutrition Programs.
- 5. Evaluation enhances program credibility. During a time of increasing need and decreasing resources, professionals and the general public question the significance of service programs. Reports of NET Program performance can enhance the credibility of the program within the State agency and the department, to policy makers as well as to program stakeholders.
- 6. Evaluation supports the NET Coordinator's professionalism. Seeing the results of one's efforts promotes a sense of self-satisfaction. Knowing that one's management skills and professionalism contributed to the success of a program increases job satisfaction. Professional managers can justify their position, show the value of their job function, and demonstrate their contribution to the organization through a well-designed evaluation of program objectives. Coordinators can demonstrate the contributions made by NET to the overall Child Nutrition Programs and to the education programs in their State.

The Evaluation Challenge

Evaluation is one of the most challenging aspects of program management. Too often, program managers avoid this crucial aspect of program management because of concerns stemming from a lack of understanding about the evaluation process itself. To learn how to plan and manage the evaluation of the NET State Plan objectives, universal concerns need to be acknowledged. Following is a discussion of some of the concerns which NET Coordinators may have.





Concern #1: "With limited funding, there is not enough money to evaluate the program." This concern is a common one in most service programs. A well-designed evaluation does not have to be an expensive undertaking. The issue of limited resources for evaluation can be resolved by including evaluation as an integral part of the design of the entire program, not isolating evaluation as the expendable component of the program when cuts have to be made. Since evaluation provides data for both decision making and program accountability, without evaluation the program cannot be expected to meet the needs of the targeted groups nor conform to program requirements over the long term. The question to be answered should be, "How should the evaluation be designed?" rather than, "Should evaluation be undertaken?"

Concern #2: "As a program planner, I don't know what information to collect." The purpose of evaluation is to determine to what extent goals and objectives have been accomplished. Thus, clearly defined objectives for the program provide the basis for the evaluation. When objectives are not clear, the program planner does not know what information to collect.

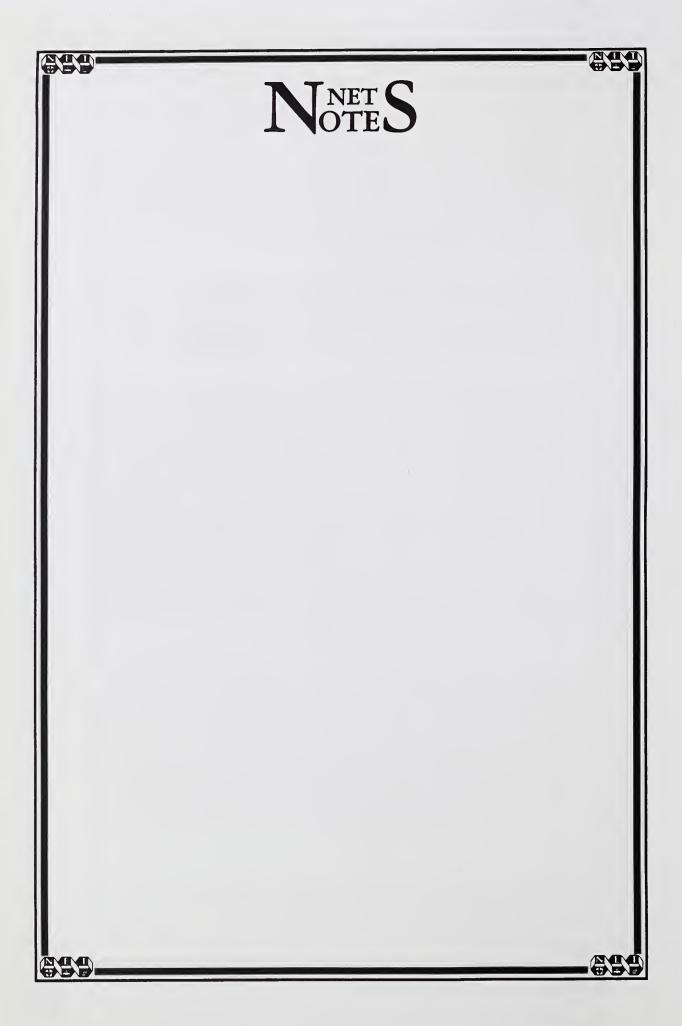
Concern #3: "There are too many variables affecting the outcomes for me to isolate the effects of the program." With any aspect of a food service training program or a nutrition education program, there are certainly many variables which affect the outcomes. For example, when training school food service managers to implement the Dietary Guidelines in their schools, a few of the variables which impact their ability to do so include: the on-going support and knowledge of the district school food service directors; the support of the school administration and faculty; the support and interest of the parents and community; the management skills of the manager; the availability of food items which are consistent with the Dietary Guidelines; and the knowledge and skills of the food service assistants. Waiting for program operations to become simpler is not reasonable. If anything, most of the work involved in the operation of Child Nutrition Programs becomes more complicated each year. Evaluation will not be as precise as that used for a controlled experimental project, but an organized approach to evaluation that acknowledges lack of control of some variables is preferable to no evaluation at all.

Concern #4: "Program evaluation may lead to information we would rather not have." Good evaluation design provides an opportunity to obtain varying perspectives of program implementation. It is certainly possible that the evaluation data may reveal difficulty in accomplishing objectives, just as it reveals success. Impact evaluation of a long-term training program may tell the evaluator that participants did not actually change their behavior on the job. Input evaluation of a nutrition education program may reveal that lack of teacher training severely limited the use of a State curriculum guide in elementary classrooms. Knowing more about the implementation and results of a program provides critical information for long-term program improvement and for maximum use of available resources.

Summary

Program evaluation is basic to the effective management of a State NET Program. Evaluation is required in the NET regulations as a measure of program accountability and is necessary for management decision making. Planning and implementing evaluation is an integral part of NET Program management.









Chapter II Developing and Managing Evaluation

Evaluation as Part of the A-P-B-I-E Process

The NET Needs Assessment Guide (1994) described NET program management as the Closed Loop Assessment-Planning-Budgeting-Implementation-Evaluation Process (A-P-B-I-E Process). This process illustrates how evaluation is an integral part of program management. Although the results of evaluation closes the loop in the management process, planning for evaluation begins when the State Plan is developed. A systematic approach to planning and implementing the evaluation aspect of the NET Program is essential to obtaining and using evaluation data to improve the Program.

NEG



The needs assessment provides the futuristic perspective for the process. As the initiating component, the primary focus of needs assessment is the determination of "where are we now in relation to where we want to go?" "What are our greatest challenges today and what challenges do we project for tomorrow?" "What needs must be addressed if we are to achieve our goals?" "What obstacles must be overcome?" "How will that impact be felt?" "What should be our priorities?" The answers to those questions provide the basis for planning, including the identification of specific objectives to be accomplished for a given period of time and the outline of the action plan which will contribute to the achievement of the objectives.

Answers also provide the basis for all subsequent resource allocations (budgeting) to be made during that time-frame.

Following *implementation* of the State Plan, the *evaluation* assumes a historical focus by returning to the same questions asked in the needs assessment, "After all our effort, where are we NOW in relation to where we want to go?" "What didn't work and why?" "What impact did we have in addressing our defined needs?" "Were our priorities appropriate relative to our current conditions and to the changes taking place in our State?" The answers to those questions, in turn, provide the basis for a re-examination or re-assessment of projected needs. And, thus, the cycle begins anew.

The assessment and evaluation components represent both the initial and final phase of a closed loop assessment-planning-budgeting-implementation-evaluation (A-P-B-I-E) process or cycle.

(NET Needs Assessment Guide, 1994, pp. 10-11)

ること

Note S

839.



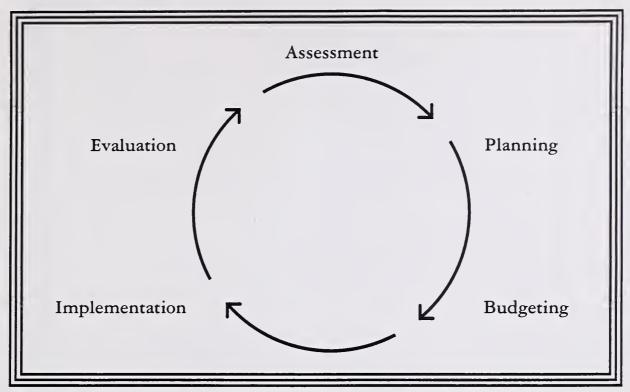


Illustration 1. Closed Loop Assessment-Planning-Budgeting-Implementation-Evaluation (A-P-B-I-E) Process. Source: NET Needs Assessment Guide (1994), p.10

Since evaluation is the final step in the Closed Loop A-P-B-I-E Process, it brings the program manager full circle, back to the questions raised in the needs assessment process when obstacles that must be overcome are identified, and needs that must be addressed to meet the goals and objectives of the NET Program are defined. Using the results from evaluation, the Coordinator can make the decision to repeat a program initiative that was successful, to develop program improvements for an initiative, or to eliminate a program initiative in favor of a different approach to meeting identified needs. Developing an evaluation plan is essential to obtaining data necessary to make program management decisions.





Planning for Evaluation

The plan for NET evaluation must:

- Be designed as part of the State Plan development process to coordinate with the objectives defined in the State Plan;
- Be easily understood and used;
- Provide for a constantly evolving, ever-changing, continuous process of evaluation;
- · Be perceived as purposeful;
- Include a variety of types of evaluation that address various aspects of planning and implementation;
- Provide a basis for improved practices and program management; and
- Promote objectivity in decision making.

Developing a plan for NET evaluation involves several important decisions. Since evaluation is a continuous process, the decisions should be reconsidered periodically. Decisions which influence the overall evaluation plan involve determining:

- A balance between the use of formative and summative evaluation,
- A balance between formal and informal evaluation design,
- A balance between the use of quantitative and qualitative methods, and
- Fair evaluation in relation to proximity of outcomes.

A balance between the use of formative and summative evaluation

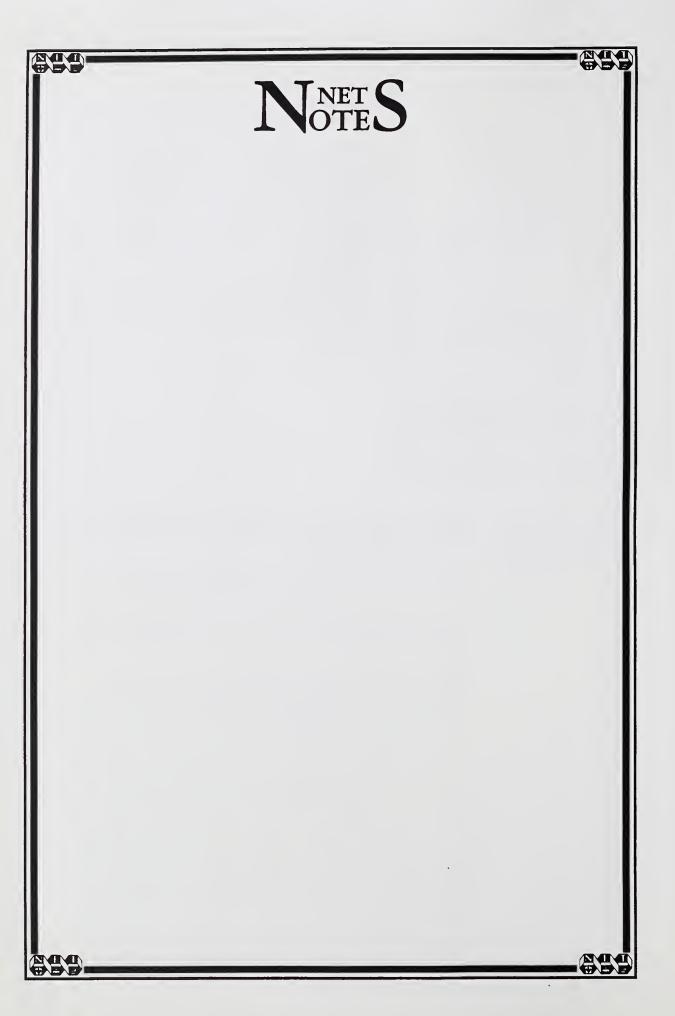
When planning the evaluation phase of the A-P-B-I-E cycle, two questions should be addressed:

"When should the design of the evaluation process begin?" and

"When should the data collection aspect of the evaluation process begin?"

There is no evidence of debate among evaluation experts as to the answer to the first question. The plan for evaluation must be developed during the overall program planning process, since evaluation design is an integral part of planning. Evaluation questions (or evaluation objectives) must parallel program objectives if evaluation is to determine the progress made by the program or project, and guide future planning decisions. Specific evaluation questions should be defined for each objective in the State Plan. For example, if the State Plan includes an objective to develop a program to positively influence the food choices made at school by middle school students, the evaluation plan must answer the question, "To what extent did the program make a positive difference?" If the State Plan addresses the preparation of K-3 teachers to select appropriate nutrition education materials and use them effectively in their instruction, the evaluation question to be







use of formative and summative evaluation (Continued)

A balance between the answered is, "How appropriately can teachers make material selections and use them effectively in their nutrition education instruction?" How data will be collected to answer these questions must be determined as the plan is developed.

> The answer to the second question, "When should the data collection aspect of the evaluation process begin?" depends on how the evaluation data will be used. The data collection should begin with the initiation of the overall program planning process when information is needed for monitoring the progress of the program and making adjustments to promote program success (formative evaluation). When evaluation information will be used for program accountability and for making planning decisions that affect the next cycle of the program, the data should be collected as the projects and activities are completed (summative evaluation).

> Consider the example of a six-week staff development program for health educators for which a new training workbook has been developed. Formative evaluation needs to be included within the overall evaluation design of the staff development program and data should be collected during the implementation phase of the program. With this approach, the NET Coordinator might find that the post-tests given at the end of sessions two and three indicate low scores among the participants. Obviously, mid-program corrections need to be made, but without the formative evaluation data provided by the post-tests, the Coordinator would not know that program adjustments were needed. With these data, the Coordinator can make decisions about what changes should be made.

> Summative evaluation should also be included as part of the evaluation design. This will involve collecting data after the health educators' staff development program has been completed. These data should be used to make decisions about whether the staff development program for health educators should be continued as is, terminated, offered to more teachers or fewer teachers, or offered to different categories of teachers. Additionally, the summative evaluation data are needed for program accountability and for justification of the expenditure of NET funds.

829 Note S



A balance between formal and informal evaluation design

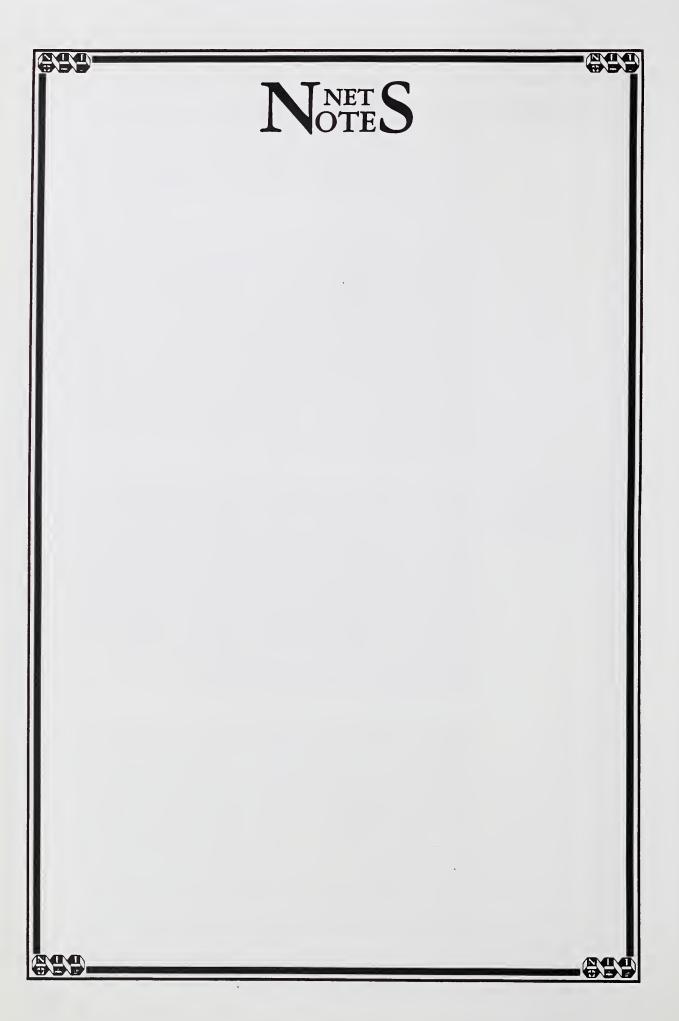
The formality of evaluation design can be placed on a continuum. Based on available time, budget, and expertise, an evaluation plan may be quite formal and highly structured. When taken to an extreme degree of sophistication, an evaluation can become so cumbersome, time-consuming, and expensive that the cost of producing the findings cannot be offset by subsequent program improvements.

At the other extreme of the formality continuum is the collection of informal information based on a "how do you feel and what do you think" five-question form given at the conclusion of a three-hour seminar. Though this information has its purpose, too often this is the sum and substance of an evaluation effort. Because the purpose of evaluation is to determine the contribution of programs and activities to the overall success of the NET Program, substantive data must be collected. If the evaluation is so "slap-stick" and informal that the results are meaningless, there will be no data base for determining program accountability nor for making program improvement decisions.

A balance between quantitative and qualitative evaluation methods

Another continuum relating to evaluation design addresses the issue of quantitative or empirical research at one end, and qualitative research on the other. Achterberg (1988) finds that, traditionally, nutrition education has focused primarily on quantitative evaluation. Quantitative evaluation seeks to determine effectiveness by use of pre- and post-assessment measures. It typically includes evaluation designs and techniques that produce discrete, quantifiable data which can then be subjected to rigorous statistical treatment. Achterberg argues that exclusive use of this methodology fails to provide the decision-maker with any data on how or why change occurred or failed to occur.

Qualitative evaluation, in contrast, focuses on an explanation of "why" and "the meaning behind" an occurrence or outcome rather than differences of statistical significance. It provides more complete information on the subtleties that contribute to differences found when comparing pre-and post-measurements. It can also provide valuable information on effects, happenings, and changes which were not anticipated in the original evaluation design, as well as information on subjective phenomena which cannot be objectively observed or measured. Qualitative evaluation uses methods such as open-ended interviews, observations, extracts from documents, and expert judgments to assess "hard to mea-





A balance between quantitative and qualitative evaluation methods

(Continued)

sure" phenomena (Wholey, 1991). One should be aware that although qualitative evaluation provides a valuable perspective of the total picture of a program or project, the analysis of qualitative data can often be labor intensive.

When considering whether to use qualitative evaluation designs in lieu of quantitative designs, Tukey states, "it is better to have an approximate answer to the *right* question, which is often vague, than an exact answer to the *urong* question, which can always be made more precise." (Lincoln and Guba, 1985). Filstead calls for the use of both designs, suggesting that the qualitative methods "provide the context of meanings in which the quantitative methods can be understood." (1979).

It is doubtful that the excessive use of either type of evaluation can provide the array of information needed to make wise decisions. Rather, Achterberg and others encourage the use of a combination of qualitative and quantitative methods "so that the flaws of one method can be compensated for by the strengths of another method." (1988).

Fair evaluation in relation to proximity of outcomes

An issue in designing evaluation for nutrition education programs involves determining fair evaluation in relation to the proximity of the outcomes of the program or project to the actual implementation (St.Pierre, 1982). For example, when a nutrition education curriculum for grades K-3 is provided to teachers throughout a State, by specifying a primary outcome of changing student's health status, the evaluator sets up the program for an unfair evaluation. Changing health status is a long-term outcome that is affected by myriad factors, not changed by a single intervention. More realistic short term results for a classroom nutrition education program could include increasing student knowledge about nutrition and possibly changing attitudes about food and eating.

The NET Coordinator may be faced with a dilemma when specifying program outcomes. Policy makers and sometimes program developers can have different expectations of the outcomes of nutrition education and training programs. The evaluator of a nutrition education program or project must balance a program practitioners' reasonable expectations of changes in knowledge and attitude toward food at school, with policy makers' expectations of more distant outcomes of changes in food-related behaviors. St.Pierre (1982) discussed proximity of an outcome and three other characteristics:



<u>849</u> Note S

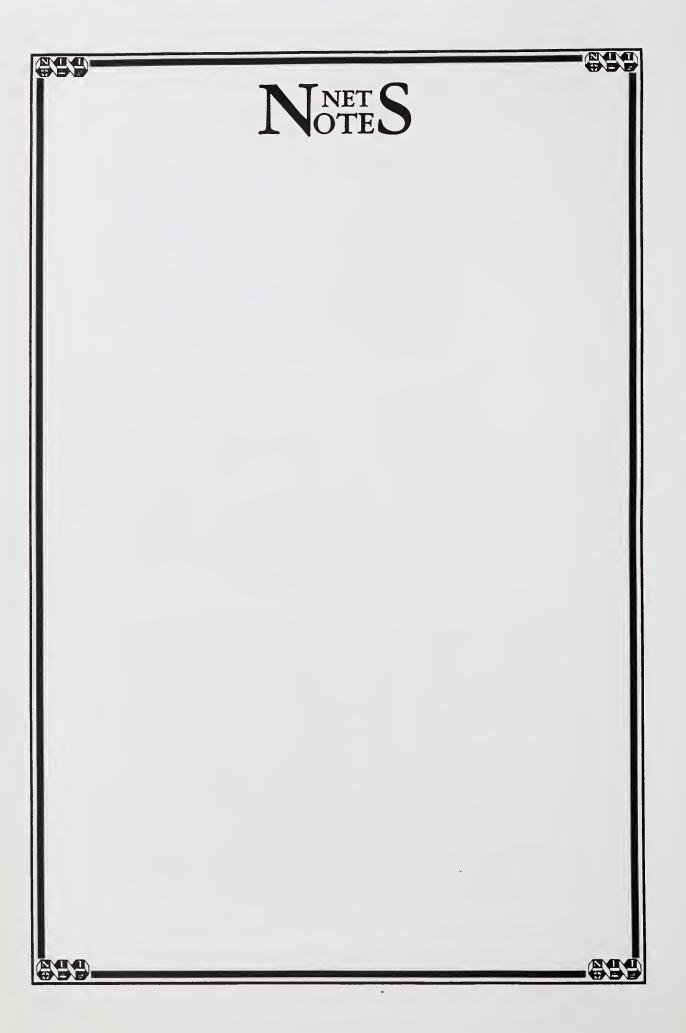


Fair evaluation in relation to proximity of outcomes
(Continued)

First, proximate outcomes (short-term results) tend to have low policy relevance and social desirability while distant outcomes (long-term results) are high on this dimension. Improvements in health status or in nutrition-related behaviors are more socially desirable outcomes and are more policy relevant than is change in knowledge about nutrition. Second, educational programs usually have a larger effect on proximate outcomes than on distant ones. It is easier for a nutrition education program to teach children nutritionrelated knowledge than to change their health status. Third, if a relationship is found between participation in a program and an outcome, it is easier to argue that the relationship is causal when the outcome is proximate rather than distant. Unfortunately, limiting evaluation to distant outcomes makes the detection of effects quite difficult; and if an effect is found, attribution to the treatment is almost impossible.

St.Pierre, one of the authors of the 1981 evaluation of the NET Program summarized the research at that time, "Little research evidence supports a clear and consistent relationship between increased nutrition-related knowledge and behavior change. Given the range of factors that influence a child's food-related behavior, behavior change may not be a fair outcome for in-school nutrition education programs which, in general, do not address many of the out-of-school factors that influence children's eating patterns." Fourteen years later, this opinion seems to still hold true and has important implications for the design of fair evaluation for nutrition education programs and projects.

The NET Coordinator should address these four important decisions discussed above regarding how an evaluation plan is designed. Decisions regarding how to approach each issue will be individual to each evaluation plan. The decision-making process and the design of the NET Evaluation Plan will be enhanced when input is obtained from individuals involved in the various projects and activities of NET and from individuals who represent participant groups. The Needs Assessment Guide For The Nutrition Education and Training Program (1994, page 38) includes a discussion of the advantages of involving a task force in the needs assessment process. If the State agency does have a Needs Assessment Task Force or some other advisory group for the program, it is desirable to obtain input from that group regarding the evaluation plan as well. The task force should be provided the report of the needs assessment and their input should be sought to define program priorities and to develop evaluation plans. The task force can also provide valuable insight on practical evaluation measures for the defined program objectives and on standards to assess acceptability or measure success for the objectives.





Managing Evaluation

Developing a plan for evaluation involves the same steps as developing a plan for a needs assessment.

Step 1. Develop an evaluation plan for each objective in the State Plan.

Include:

- Evaluation questions for each objective (program, project, or performance objective),
- For each evaluation question, identification of data to be collected,
- For each evaluation question, identification of standard to assess acceptability,
- For each evaluation question, identification of appropriate data collection method, or technique, and
- A work schedule with assigned responsibilities.
- Step 2. Determine how the evaluation plan will be managed, either in-house or through a contractor.
- Step 3. Monitor the evaluation plan.
- Step 4. Analyze and interpret results of the evaluation.
- Step 5. Use the evaluation findings in management decision making, determining program policies and procedures, and for program accountability.









Step 1. Develop an evaluation plan for each objective in the State Plan.

The State Plan objectives must be written in such a way that they can be measured. Then, the NET Coordinator should develop an evaluation plan for each objective in the State Plan. This should occur during the developmental phase of the State Plan.

When writing objectives for a specific program or project, the same process should be used. The program, project, or performance objectives, define what should be evaluated.

The evaluation plan for each objective should include the following information:

Evaluation
question(s) for each
objective
(program, project, or
performance).

Evaluation questions identify the specific purpose(s) of the evaluation, that is, what you want to know from the evaluation. The evaluation questions should be developed during the planning phase of the program, not after the program has begun, or has been completed. They are essentially the reasons for evaluation, and as such, affect the baseline data that will be collected and the data collection method(s) chosen. Some evaluators write evaluation questions as evaluation objectives using the traditional style of objectives. For purposes of NET evaluation, evaluation questions are recommended since they provide a clear focus for the subsequent collection of evaluation data and are not confused with State Plan objectives.

The evaluation questions should address all components of program or project planning and implementation. The questions should be written to obtain a multi-dimensional view of what is working or not working as planned, what worked or did not work as planned, and why. This brings up the issue of types of evaluation and evaluation models. An explanation of one of the most popular evaluation models is included in Chapter 3. The model includes five types of evaluation which provide a comprehensive approach for defining evaluation questions.

Some examples of evaluation questions of various types are shown below.

- How did the situation in which the program was developed and implemented affect the program implementation?
- Were the resources for the program adequate for program implementation?
- ₩as the program delivered as it was planned?
- What processes during program implementation were successful?
- ₲₲ To what extent were program objectives met?
- ₩ What changes occurred as a result of the program?





For each evaluation question, identify data to be collected.

The evaluation questions are used to determine the types of data that must be collected. Some examples of data include: attitudes, behavior, opinions, numbers, program costs, productivity, time, behavior, quality, personal reactions, perceptions, or observations.

Data to be collected for each evaluation question should be described in the evaluation plan. The data descriptors should meet the following criteria:

- 1. Be meaningful and relevant to the evaluation question (or evaluation objective);
- 2. Be available and easily accessible;
- 3. Provide value to justify the cost of obtaining the data. In other words, the cost of the evaluation should be appropriate to the cost of the program;
- 4. Be observable and measurable; and
- 5. Be timely.

Data collection may involve an entire NET-targeted population or it may involve a sample. When quantitative evaluation methods are used, the sampling protocol becomes very important. When qualitative methods are used, sampling becomes less of an issue. There are a variety of sampling procedures. A description of probability sampling is included in Appendix E. This Sampling Protocol was originally prepared for use by NET Coordinators in sampling for Statewide needs assessments. Additional references on sampling are listed in the Selected Bibliography of the Sampling Protocol.

For each evaluation question, identify a standard to assess acceptability.

Standards represent the pre-determined yardsticks that will be used to measure success or to determine acceptability. Standards for evaluation must be

₫ reasonable,

△ measurable, and

△ observable.

They may be described with numbers, percentages, or behavioral descriptions that are clear to everyone involved. As part of the planning process, the NET Coordinator, in collaboration with others involved in the program, should set reasonable standards using his/her expertise in the subject area.

It is most important that the standards be defined during the



NoteS

629



planning stage with input from those people who will be involved in the program or project. For example, in the case of training programs, the NET Coordinator, other State staff members involved in the training, representatives of the target groups, and the trainers themselves should have input into what will be considered "acceptable" results of the training.

Consider the example of a training program for school nutrition managers on using production records to forecast amounts to prepare for a choice school lunch program. The performance objective might be stated as,

"The manager will demonstrate the efficient use of production records for forecasting by reducing the amount of leftovers of main dish items prepared in the school kitchen."

The evaluation question could then be, "Can the manager use production records for forecasting to reduce the amount of leftovers of main dish items prepared in the school kitchen?" The data for this evaluation question could be obtained at the end of the training program by having the managers complete a case study in which they demonstrate use of production records for forecasting where they reduce the potential amount of leftovers. The completed case study could be included in a portfolio assembled to reveal the manager's proficiency in content areas addressed in the training program. The standard, which was pre-determined when the evaluation was planned, could specify that 85% of the managers in the training program could use the production record for forecasting based on defined criteria in the case study.

For each evaluation question, identify an appropriate data collection method or technique.

The evaluator is now ready to identify a data collection method or technique (Phillips, 1991). When selecting data collection methods and techniques, consider the following:

data needed to answer the evaluation question.

For example, when data are needed to determine the extent to which training program participants have attained specific skills defined by performance objectives, the data collection method must provide evidence of the new behavior. A method which only obtained data relating to participants' personal reactions to the training would not answer the question, "can they?"

1. Whether the data collection method can be used to attain the

NoteS



For each evaluation question, identify an appropriate data collection method or technique.

(Continued)

2. How the evaluation data will be used.

Knowing the purpose of evaluation data will help to guide the selection of the method used to obtain it. An effective evaluation gathers facts. The object is to determine which facts are needed to answer the evaluation question.

3. How the data will be collected, analyzed, and reported. The evaluator must consider how the data will be collected, tabulated, summarized, analyzed, and reported. The time involved and the expertise needed to collect the data will influence the data collection method chosen. When external consultants with expertise in research design are given responsibility for evaluation, a more sophisticated and indepth evaluation method can be used. Consultants generally work on a contractual basis and are able to devote more time to managing a comprehensive evaluation than can a State agency staff member.

4. Who will use the information.

The target audience for the evaluation summary or report should be considered when determining the methods that will be used. If the target audience is policy makers at the State or national level, different evaluation questions may be defined than when the target audience is instructional designers for a training program.

The Needs Assessment Guide For The Nutrition Education and Training Program (1994, page 44) includes a detailed description of the five most commonly used data collection methods for needs assessment studies: (1) written surveys, (2) interviews, (3) observations, (4) data reviews, and (5) focus groups. Since these same methods are often used for evaluation, the evaluator is encouraged to review each technique (method) before deciding which is the most appropriate to obtain data for answering the evaluation question while conforming to the constraints of the program.

In addition to the methods already mentioned, the description of three data collection methods useful for evaluation have been included in the Appendices of this Guide. These include: the case study, the participant reaction questionnaire, and the portfolio assessment.



Note S



For each evaluation question, identify an appropriate data collection method or technique.

(Continued)

A work schedule with assigned responsibilities.

Other data collection methods which are appropriate for use with NET evaluation include longitudinal studies, experimental design, and attitude surveys. Because these methods require considerable time and expertise to design and administer, a description has not been included in this Guide. However, most research methods textbooks include a detailed description of these procedures. The Reference section includes a listing of several such textbooks.

Regardless of the data collection method to be used, the evaluator should develop an instrument that is consistent with good design principles. Refer to Appendix A for Guidelines for Developing Evaluation Instruments.

Managing program evaluation requires a systematic approach based on a management plan. The evaluation plan should include:

1. A work schedule. The work schedule should identify the major activities or tasks required to answer the evaluation questions or implement the data collection method. As with any aspect of program management, determining the time necessary to complete an activity is often, at best, guesswork. However, it is important to designate beginning and ending times for each activity leading to a completion time-line for each part of the evaluation.

Two of the most popular project management techniques are network models and Gantt charts (Martin and Trumbley, 1987). One of the most commonly used network models is the Program Evaluation and Review Technique (PERT). The PERT diagram is useful in the planning process for showing interdependent and dependent tasks. This can be a major aid to the program planner by helping to identify which tasks can be performed simultaneously and which tasks must be performed consecutively. The program planner is required to think backwards from the desired outcome, identifying the key tasks that must be performed to achieve the evaluation objective. A flow chart is then planned that indicates the path or route for performing those tasks with estimated time-lines for completion. The critical path is identified as the path which links the tasks that must be performed consecutively, and therefore, take the longest time to complete.



Note S



If that sequence of tasks is not completed on time, all other aspects of the project are negatively affected.

A work schedule and assigned responsibilities.

(Continued)

The Gantt chart is based on a horizontal time scale that is easily referenced for quick-assessment of project status and is particularly useful for project monitoring. To develop a Gantt chart, the program planner lists tasks or activities down the left vertical axis of a graph or chart. With time intervals on the horizontal axis, the beginning and ending time for each activity is projected. The time interval may be days, weeks, or months. Time frames for the estimated performance and completion of specific tasks are reflected with the use of solid lines under the date or time interval shown on the graph. Actual times required for the completion of the tasks are then shown with a broken line directly under the solid project time line. This management tool is easy to develop, is useful for identifying overlapping activities, and serves to keep a project on schedule. Gantt charts and PERT charts or diagrams are described in the Needs Assessment Guide For The Nutrition Education And Training Program (1994, pp. 46-48).

2. Identification of personnel assignments. When evaluation is conducted "in-house," the NET Coordinator may be responsible for all of the evaluation tasks. If evaluation is part of a project contract or is a separate contract, the NET Coordinator may serve as the project director or the project monitor. In either case, it is important that for each activity or task, the responsibility is designated for someone. Neither Gantt charts or PERT diagrams include personnel assignments. However, a Gantt chart can be adapted to include the person responsible for each task, in a column parallel to the task column.

An example of a systematic plan for evaluation of NET State Plan objectives is included in Chapter 4.

SED Note S 628 829



Step 2. Determine how the evaluation plan will be managed, in-house or through a contractor. For contracted projects, provide funding for the contractor to include an evaluation as part of the project.

When a project proposal includes a description of evaluation, make sure the work schedule includes the tasks, completion dates, and personnel responsibilities for all aspects of the evaluation plan. This will provide the NET Coordinator a basis for program monitoring throughout the evaluation process. Section 6 of the NET Needs Assessment Guide (1994, pp. 56-63) provides guidance for developing requests for proposals and managing external projects.

Step 3. Monitor the evaluation plan as part of monitoring the implementation of the NET State Plan.

A detailed evaluation plan for each State Plan objective enables the NET Coordinator to monitor implementation. A Gantt chart provides a visible tool to monitor the beginning and completion of each task of the evaluation plan.

Step 4. Analyze and interpret results of the evaluation.

After data have been collected, the evaluator should follow some basic guidelines for managing the data. (Phillips, 1991).

- Review the data and eliminate data collection instruments which have been completed incorrectly or which are incomplete.
- Use all the relevant data. Of course, any program manager desires positive evaluations and this leads to a built-in bias. Even though data will likely contain both positive and negative information, all of it should be analyzed.
- Maintain confidentiality of data. The same security used for maintaining confidentiality during data collection should be continued during analysis and reporting.
- Use the simplest analysis possible to answer the question. In most cases, descriptive analysis can be used to answer the questions posed in NET evaluation. Descriptive treatment of data simply describes the data. The results can be presented as means, medians, modes, and frequency distributions. This analysis provides data that are easy to display in various kinds of charts. However, it cannot be used to identify correlations or to make inferences about a large population from a small sample.

If inferential treatment is necessary, the appropriate statistical tests should be used. This type of statistical treatment of data describes cause and effect connections or relationships among variables. Use caution when designing the analysis so that the resulting data do not become so burdensome and detailed that the reader of the report cannot determine the key results. Below is a brief summary of some common tests used in inferential statistics.





Inferential Treatment of Data

Statistical inference comprises those methods concerned with the analysis of a subset of data leading to predictions or inferences about the entire set of data. Some of the more common tests used in inferential statistics are identified below.

<u>Test</u>	<u>Definition</u>
t-test	To compare the means of two groups. If the means are far enough apart, the t-test will show a significant difference, indicating that there is a statistical difference between groups.
ANOVA (one-way)	To simultaneously compare the means of more than two groups.
Pearson Product Moment Correlation Coefficient (r)	Shows whether there is a correlation or relationship between two sets of numbers. <i>r</i> does not imply a cause-effect relationship among variables.
Phi Coefficient	A nonparametric correlation statistic where both variables are divided into two parts.
Spearman Rank Order Coefficient (<i>Rho</i>)	A correlation statistic where the data are ordinal in level, and where the data can be placed in rank order. The highest score is ranked "1" and so on.

Illustration 2. Common statistical tests.

Adapted from: American Society for Training and Development. (1991). Statistics for HRD practice. INFO-Line. Author. p.10.

The data analysis provides the evaluator with the basis for developing the evaluation report. The type of report needed will vary with the program or project, and the evaluation questions. If the evaluation is needed only by the NET Coordinator for making management decisions, it may not be necessary to prepare a formal report; rather the Coordinator may draw conclusions from the analysis and act upon them. An example of this situation would be the formative evaluation of a training program where the results will be used immediately to make the changes needed for the next scheduled training program.



839 NoteS

(



In most situations, the evaluation results will be used both for making management decisions and for program accountability, thus a formal report will be needed. Evaluation should be designed for each program objective included in the State Plan, and the results should be provided with the subsequent year's State Plan. When preparing an evaluation report, use tables and graphs to display the data, limiting narrative to a description of the visual displays. More information on how to write a report to present the findings can be reviewed in the NET Needs Assessment Guide (1994, pages 52-53).

For evaluation to be truly meaningful, a review of the analysis should include time for reflection in an attempt to determine the real meaning of the results. Reflection upon the evaluation findings along with personal observations and impressions, can provide the evaluator with valuable insight to answer the evaluation questions (Steele, 1991).

Evaluation becomes most valuable when it helps the program manager think creatively and think ahead. Evaluation by its very nature deals with the recent past. The program manager must analyze current trends, changes in the program, and changes in the environment, along with the evaluation results to move the program forward to fit the future. Having examined the analysis with a creative and critical eye, the report can be developed.

Step 5. Use the evaluation findings for management decision making, for determining program policies and procedures, and for program accountability.

The purpose of evaluation is to provide data to continuously improve the NET Program and justify its funding. Thus, the gathering of evaluation data and analyses will be of no value unless the findings are used for continuous program improvement.

Summary

Developing and managing evaluation for the NET Program involves a step-by-step process. The process includes the development of an evaluation plan, management of the plan, monitoring the plan as part of the NET State Plan, analyzing and interpreting the results of the evaluation, and using the evaluation findings for program management. The evaluation plan should include evaluation questions for each State Plan objective. For each evaluation question, the NET Coordinator should identify: the data to be collected, a standard to assess acceptability, a procedure including an appropriate data collection method, and a work schedule with assigned personnel responsibilities. Evaluation closes the loop in the Assessment-Planning-Budgeting-Implementation-Evaluation Process for managing a State NET, a project, or program.

Note S





Chapter III THE CIPPI Model for Evaluation

Characteristics of an Evaluation Model

C

A State NET Program is a dynamic program continually changing to meet the identified needs of the targeted populations. Thus, NET is essentially developmental in nature. A comprehensive evaluation model is needed to examine the variety of projects and activities that are part of NET. An evaluation model appropriate for NET should:

- I
- recognize the dynamic nature of the Program;
- address the interdependence between contextual preconditions, program delivery attributes, and program outcomes (Edwards, 1986); and
- provide the State NET Coordinator with the evaluation information necessary to make management decisions and conform to the regulatory requirements of the NET Program.

P

There are a variety of different models for evaluation. The model presented in this Guide was selected because it addresses basic elements appropriate to the evaluation of programs which are developmental in nature. The CIPPI Model for evaluation is a multi-dimensional approach for the collection and assessment of data needed for making effective and cost-efficient program management decisions. This evaluation model precludes the evaluator from taking a myopic view of the scope of evaluation or a piecemeal approach to evaluation design. The CIPPI Model is one of the most commonly used frameworks in the design of a comprehensive evaluation plan. The original model for CIPP evaluation was based on the work of Stufflebeam (1971) and was further refined by Trapnell (1984) to include impact evaluation, adding the I, thus becoming the CIPPI Model. This evaluation model provides both proactive information support for management decision making and retroactive information support for accountability. Thus, it is particularly appropriate for NET evaluation.

I

829 NoteS



The CIPPI Model

The CIPPI Model includes five major types of evaluation.

- C Context Evaluation
- I Input Evaluation
- P Process Evaluation
- P Product Evaluation
- I Impact Evaluation

This multi-dimensional approach to evaluation provides a way for the NET Coordinator to examine NET as a total phenomenon, rather than a series of isolated projects and activities. Each type of evaluation provides a framework for determining evaluation questions. Review an example of a NET Evaluation Plan, shown in Chapter 4, to see examples of questions for each type of evaluation.

The following is a brief description of each type of evaluation comprising the CIPPI Model. Chapters 5, 6, 7, 8, and 9 provide a detailed description of each type of evaluation.

© CONTEXT EVALUATION (Chapter 5)

This type of evaluation focuses on the overall social, political, economic, and environmental setting in which NET is planned and implemented. It can be considered a situation evaluation. It addresses the general context in which the Program is implemented and how this context affects the achievement of the Program's goals and objectives. Success or the perceived lack of effectiveness of a State NET Program may lie in the context in which that program operates rather than in the design or delivery of the program itself. Data obtained for context evaluation may come from needs assessment. It describes problems, organizational climate, barriers, opportunities, and value orientations. This type of evaluation is generally conducted with administrative or management personnel in the State agency, school districts, sponsoring agencies, or cooperating organizations.

♣ INPUT EVALUATION (Chapter 6)

Input evaluation assesses the appropriateness and adequacy of all human, fiscal, and physical resources that go into the design and implementation of a given program or initiative. Included are such resources or inputs as equipment, facilities, materials, supplies, curriculum guides, entry-level ability of students, and the competency level of the instructional staff providing the education or training program. Input evaluation is usually conducted with the individuals who are involved in planning and implementing a program. This might include program administrative or management personnel in the State agency, school districts, or sponsoring agencies; or trainers and others involved in implementing a program or project at the local level. Program participants could also be targeted for input evaluation.







In a training program, input evaluation would provide information about the adequacy and appropriateness of the training materials. Included would be the identification of performance objectives, the design of the instructional activities, the planned sequence of the instructional activities, and the types of training materials to be used by both the instructor and participant (or student). Furthermore, it would address the adequacy of the equipment and facilities, the entry-level behaviors of the participants, and the competency level of the instructor.

Process evaluation provides the NET Coordinator information about the extent to which the program is implemented as it was designed; whether resources are used appropriately during the implementation phase of the program, project, or initiative; and, if the students are meeting interim performance objectives. This type of evaluation addresses the guiding management principles of "plan your work," and "work your plan." It provides information on whether and/or how well the program or project plan was implemented. Formative evaluation, which occurs during the implementation phase of a program or project, is especially critical in process evaluation, as it can provide an array of data for making changes during the implementation of a program or project. Individuals who are usually the audience for this type of evaluation include program administrative or management personnel in the State agency, school districts, and sponsoring agencies; trainers and others involved in implementing a program or project at the local level; and program participants.

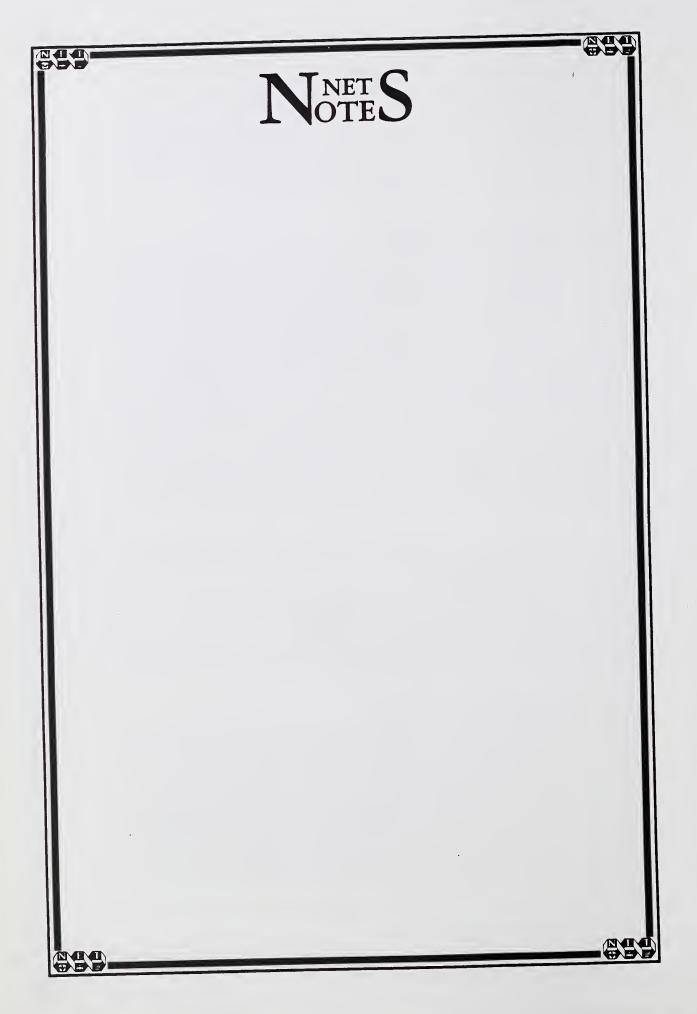
Process evaluation, which occurs during implementation of a training program, can provide feedback regarding:

- structor in utilizing the resources,
- interim performance gains of the participants (or students), and
- see implementation of the program as it was designed.

Using the evaluation data, the NET Coordinator can make important changes in the program during the implementation process.

☑ PRODUCT EVALUATION (Chapter 8)

This is the type of evaluation with which most people are familiar. It focuses on participant (or student) outcomes, ability levels, and performance after the completion of a program. It may be referred to as the "proof of the pudding" because it determines the extent to which students attain the performance objectives included in the curriculum guide or training workbook and whether the students can apply those skills in a variety of settings. Product evaluation, thus, closes the loop in curriculum and instructional design. Product evaluation audiences are primarily program participants who may include adult learners or K-12 students.





Summative evaluation is the most common design in product evaluation, as the evaluation takes place after students have completed a course or program. Although product evaluation has been the focus of the majority of evaluation efforts over the years, it represents only the tip of the iceberg in the design of a comprehensive evaluation effort.

* IMPACT EVALUATION (Chapter 9)

Impact evaluation focuses on the cumulative difference NET makes on the target groups it is intended to serve, and on the problems or needs it is intended to address. Thus, impact evaluation may be considered "the bottom line" evaluation. This type of evaluation, more than any other, provides the justification for budgetary requests, especially for budget increases, as it documents the ultimate effectiveness of the NET Program. Impact evaluation provides a picture of long-term program impact or outcome rather than the difference one training program or one nutrition education project made.

The CIPPI Model for evaluation provides the NET Coordinator a comprehensive approach to collecting and analyzing data about NET throughout the planning and implementation process. The various types of evaluation, taken together, can provide a overall picture of the program and its major projects that will support sound management decisions.

Summary

The CIPPI Model can be used to design the evaluation for large programs such as a State NET Program, large projects such as a State training program, or small projects such as a mini-grant to a school district or Child and Adult Care Food Program (CACFP) center sponsor. It provides a multi-dimensional approach to evaluation which enables the NET Coordinator to design a mosaic with each piece of collected data assembled into a whole that is more meaningful to program improvement than the individual evaluation parts.

NetS





Chapter IV Application of the CIPPI Model NET Evaluation Plan

The CIPPI Model for evaluation addresses the program performance, typical of most all evaluation models, as well as the processes involved in program delivery. By examining a program using this multi-dimensional approach, the program planner can achieve a better understanding of what makes a program successful or what needs to be improved to achieve success. Using the various types of evaluation in the CIPPI Model, the NET Coordinator should develop an evaluation plan for each NET State Plan objective.

Because context evaluation provides an analysis of the situation in which a program operates, in some cases it may be more appropriate to gather and analyze context evaluation data regarding the whole NET Program rather than specific objectives of the State Plan. This will provide the NET Coordinator insight into the social, political, economic, and physical aspects of the environment which can affect the achievement of the program's goals and objectives. For comprehensive State-initiatives, where there are several State Plan objectives with one focus, the NET Coordinator may wish to obtain context evaluation data in order to achieve an understanding of situational factors which contribute to overall program success or hinder program implementation. For such an initiative, context evaluation data should be obtained from State-level administrators and program personnel; from individuals in targeted school districts, schools, centers; and from other targeted agencies. An example could be the implementation of the Dietary Guidelines, where there are many different programs and projects with the same focus.

Impact evaluation is an examination of cumulative differences resulting from program initiatives over the long term. As such, this type of evaluation is not conducted for short-term training programs or for individual nutrition education interventions. Rather, it is appropriate to examine the impact of all aspects of training over a period of time and is appropriate to examine the difference made by an extensive training program, such as one which provides statewide training for all school nutrition managers in a 30-hour course with accompanying school district follow-up. This type of training can be expected to impact on the delivery of services in school nutrition departments. A one-time training program of six hours for school nutrition managers may affect attitudes and increase knowledge and skills, but is less likely to have a significant bottom-line impact on the delivery of food services. Impact evaluation is useful for examining the impact of several different programs or projects with the same focus, for example, implementation of the Dietary Guidelines.

The evaluation plan included in this chapter illustrates how the NET Coordinator can plan evaluation for State Plan objectives. It is provided as an example of the application of the CIPPI Model and is not intended as a requirement, or as the one recommended way to use the Model. For purposes of illustration, three different NET Strategic Plan goals have been used in this scenario



CE Note S 629



together with a State objective for each goal. For each State objective, one or more implementation activities have been included. Generally, a State would include more than one State Program objective for each goal. The CIPPI Model is applicable to projects and other initiatives as well.

Example - Plan for Evaluation for

Some Objectives in a NET State Plan

NET Strategic Plan - Goal 1: Increase to at least 75% the proportion of the Nation's schools that provide nutrition education as part of health education, as part of an interdisciplinary nutrition education curriculum, or through other means, to children and parents by the year 2000, as documented by data submitted to FNS or by other sampling methods.

State Plan Objective 1: To reach 30% of the K-4 students with nutrition education through the use of the State Wellness Curriculum, nutrition education section.

NET Strategic Plan - Goal 3: Increase to at least 50% the proportion of food service personnel/providers in Child Nutrition Programs who have received nutrition information and/or training (including food service training) by the year 2000, as documented by data submitted to FNS or by other sampling methods.

State Plan Objective 2: Develop and implement a train-the-trainer program to present a 10-hour Dietary Guidelines workshop to at least 400 managers.

NET Strategic Plan - Goal 2: Increase to at least 50% the proportion of child care and summer program facilities/sites that provide nutrition information/education to children, caregivers, or parents by the year 2000, as documented by data submitted to FNS or by other sampling methods.

State Plan Objective 3: To reach at least 1000 children in child care centers with nutrition education learning experiences consistent with the principles of the Food Guide Pyramid.





Plan for Evaluation

State Plan Objective 1: To reach 30% of the K-4 students with nutrition education through the use of the State Wellness Curriculum, nutrition education section.

Implementation Activity: Provide a training video for K-4 teachers to all school-district elementary curriculum directors showing ways to use the nutrition education section of the State Wellness Curriculum Guide.

Evaluation Questions	Data to be Collected	Standard	Data Collection Method	Work Schedule (Completion Date)	Responsibility
To what extent did district curriculum directors distribute the training video to K-4 teachers? (INPUT)	Numbers of teachers who had access to the video.	80% of the curriculum directors provided access to K-4 teachers in their school district.	Written survey to all curriculum directors	FY-98-April	NET Coordinator
What was the level of utility of the video in addressing the use of the curriculum guide in the classroom. (INPUT)	Teacher perceptions of adequacy of coverage, appropriateness of examples of classroom activities, relevance to their classroom use.	Teachers indicate overall positive utility of the video.	Focus group of K-4 teachers in 3 areas of State	FY-98-April	NET Coordinator
Are teachers using the training video? (PROCESS)	Amount of time, large group, small group, as a reference.	80% of reachers report using the video.	Survey questionnaire and focus groups-teachers	FY-98-April	NET Coordinator
In what ways did the teachers who viewed the training video use the murition education section of the State Wellness Curriculum Guide in their classrooms?	Health classes, whole-language plans, integrated with specific content areas, in school nutrition.	Teachers report a variety of uses consistent with the video examples.	Focus groups-teachers	FY-98-April and Contractor	NET Coordinator and
Is there a difference in the use of the State Wellness Curriculum Guide in classrooms where the reacher viewed the training video and those where the reacher did not? (PRODUCE)	Differences in student outcomes- knowledge, food selection skills, and attitude about food and eat- ing.	There is a difference.	Experimental design-K-4 teachers	FY-98-April	Contractor
How many K-4 students have been reached in classrooms where teachers viewed the video? (PRODUCT)	Number of students in class- room reached by teachers who reported viewing the video.	30% of students.	Survey questionnaire-teachers	FY-98-April	NET Coordinator



Plan for Evaluation

1)

State Plan Objective 2: To promote the implementation of the Dietary Guidelines in schools through a Dietary Guidelines workshop which reaches at least 400 school nutrition managers.

Implementation Activity 1: Develop a 10-hour workshop for school nutrition managers on implementation of the Dietary Guidelines. Implementation Activity 2: Train a cadre of 25 selected trainers to present the workshop in 15 locations around the State.

Evaluation Questions	Data to be Collected	Standard	Data Collection Method	Work Schedule (Completion Date)	Responsibility
What was the entry-level of the managers? (INPUT)	Highest grade level completed, previous workshops.	95% will have completed HS. 80% will have completed the menu planning workshop.	Review of registration information.	FY-98-October	NET Coordinator
2. How adequate were the facilities and materials used in the manager workshop? (INPUT)	Checklist of requirements.	95% of facilities and materials met requirements.	Instructor survey	FY-98-October	NET Coordinator
3. Did the instructional design of the workshop meet standards for best practices for adult learners? (INPUT)	Checklist - using principles of adult learning.	Instructional design will reflect best practices in 9 of 10 instances.	Expert review panel	FY-98-October	NET Coordinator
4. To what extent did the instructors demonstrate instructional competencies that motivate adult learners? (INPUT)	Observations - demonstration of expertise, empathy, excitement, and clarity.	All instructors will score a minimum of 85% on all of the observation forms.	Trained observers during the training	FY-98- (Summer)	Contractor
5. To what extent did the locations of the course reach the maximum number of managers for the designated regions of the State? (INPUT)	Number of managers who attended/are attending the workshop by location in the State, compared to maximum number expected.	75% of the managers who could attend, do attend to equal at least 400 managers.	Analysis of workshop registra- tion data	FY-98	NET Coordinator
6. To what extent are the instructors presenting the 10-hour workshop as it was designed? (PROCESS)	Instructional sequence, objectives, experiential activities, audio-visuals, process and product evaluation.	90% of the instruction is being used as designed.	Observations during workshop	FY-98- (Summer)	Contractor
7. How effective are the instruc- tional strategies of the instruc- tors? (PROCESS)	Small groups, Large group, variation, meering individual learners needs, questioning, attentiveness.	90% of the instructional strate- gies used were determined as effective.	Observations during the Dietary Guidelines workshop for man- agers	FY-98- (Summer)	Contractor

Note S 829

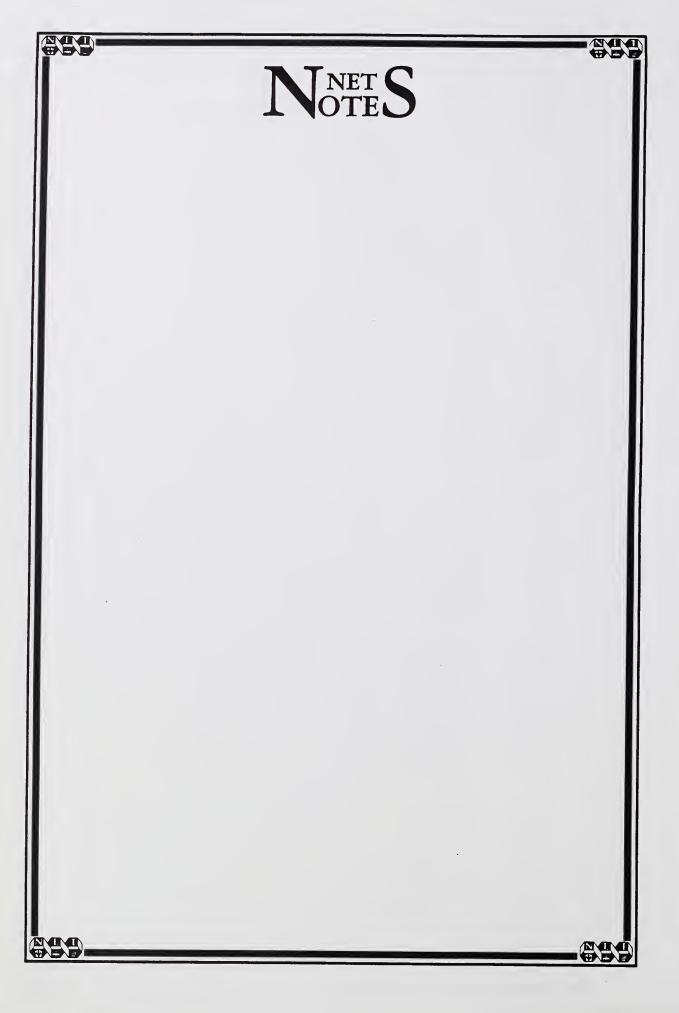


Plan for Evaluation

State Plan Objective 2: To promote the implementation of the Dietary Guidelines in schools through a Dietary Guidelines workshop which reaches at least 400 school nutrition managers.

Implementation Activity 1: Develop a 10-hour workshop for school nutrition managers on implementation of the Dietary Guidelines. Implementation Activity 2: Train a cadre of 25 selected trainers to present the workshop in 15 locations around the State.

	Responsibility	NET Coordinator and review panel of workshop instructors	Workshop instructors
The state of the s	Work Schedule (Completion Date)	FY-98	FY-98 (Summer)
are workshop in 12 wear	Data Gollection Method	Expert review of participation FY-98 authentic assessment portfolios.	Participant reaction form
amosad or summa	Standard	85% of the managers can plan specific actions.	90% of the managers respond positively.
4: Hall a cault of 27 30	Data to be Collected	Menus, modification of recipes, portion control, merchandising, marketing, parent and faculty discussions.	Perceptions that they are pre- pared to plan menus, modify recipes, merchandise and market foods that are consistent with Dietary Guidelines.
In demendation menon 1 2. Hall a carle of 2) selected trainers to present the workshop in 1) received the certes.	Evaluation Questions	8. At the end of the workshop, can the managers plan specific actions for implementing the Dietary Guidelines in their school? (PRODUCT)	9. How do the managers who completed the workshop perceive the helpfulness of it in preparing for school implementation of the Dietary Guidelines? (PRODUCT)



Plan for Evaluation

State Plan Objective 3: To reach at least 1000 children in child care centers with nutrition education learning experiences consistent with the principles of the Food Guide Pyramid.

Implementation Activity: Make available mini-grants to Child Care Center Sponsors with multiple centers who submit proposals which meet stated criteria for nutrition education projects using State-provided child care materials which focus on teaching children good food habits consistent with the principles in the food Guide Pyramid.

Evaluation Questions	Оми to be Collected	Standard	Data Collection Method	Work Schedule (Completion Date)	Responsibility
1. To what extrent does the REP include appropriate and adequate guidance for writing a proposal that can be monitored by the State Agency? (INPUT)	Measurable objectives, task ist/time frame, budget by line item, assigned responsibilities, plan for education.	All appropriate items included.	Expert review (Formative)	FY-97 - January	NET Coordinator
For each proposal, was the funding adequate for actual implementation of the project as described? (INPUT)	Budget line items were adequate to support resources needed.	90% of line irems were adequate.	Expert review panel of Spousors and State consultants. (Sammarive)	FY-98 January	NET Coordinator
Were the qualifications of the person responsible for the grant appropriate for the assigned responsibilities? (INPUT)	Job specifications, project rask list, project evaluations, individ- nals' qualifications.	90% идгетет.	Expert review	FY-98 January	NET Coordinator
How effectively did the sponsors that received grant funds implement the proposal? (PROCESS)	Time-lines of task completion; use of fonds as described in proposal; positive feedback from school administrators, faculty, syndents.	85% of items are positive.	Expert review	FY-98	NET Coordinator
5. For each mini-grant, to what extent did the sponsor accomplish the objectives stated in the proposal? (PRODUCT)	Mini grant objectives.	95% of objectives were completed,	Expert review	FX-98	NET Coordinator
How many children were reached drough the mini-grant? (PRODUCT)	Number of children in centers which completed the grant activities.	1000 children.	Mini-grant final reports	FY-98	NET Coordinator



Note S



Plan for Evaluation

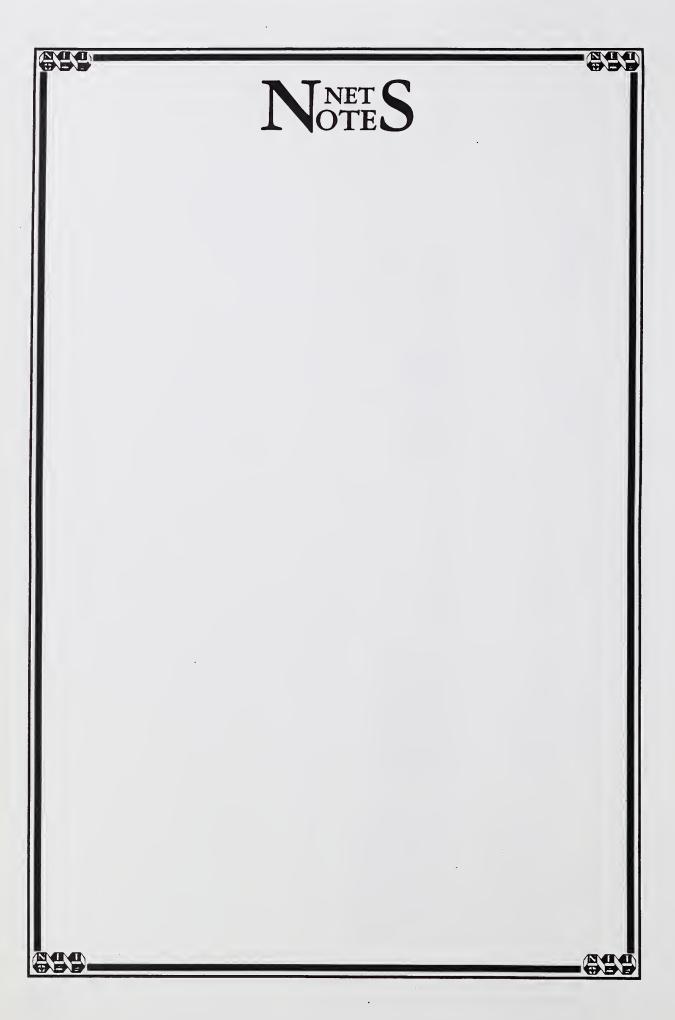
1

State Plan Objective 1: To reach 30% of the K-4 students with nutrition education through the use of the State Wellness Curriculum, nutrition education section.

State Plan Objective 2: To promote the implementation of the Dietary Guidelines in schools through a Dietary Guidelines workshop which reaches at least 400 managers.

Responsibility	NET Coordinator	NET Coordinator, CNP Staff,
Work Schedule	(Completion Date) FY-98	FY-98 and 99
Data Collection	Survey of school administrators.	Written surveys to principals, school nutrition managers, and teachers, review of production records.
Standard	75% of principals have a positive attitude.	75% of descriptors indicate positive use.
Data to be	Attitudes and opinions about nutrition education and school nutrition programs.	Menus reflect DG, classroom nutrition education, teachers and SFS manager work as a team, students' food selection decisions reflect increased awareness of the Dietary Guidelines.
Evaluation	• How did the school administrators' values about nutrition affect the implementation of the Dietaty Guidelines in schools where K-4 teachers used the video and the manager was trained? (CONTEXT)	• To what extent did the program funds used for nutrition education and training in the area of implementation of the Dietary Guidelines impact schools where K-4 teachers used the video and the manager was trained? (IMPACT)

Evaluation Guide for the Nutrition Education and Training Program







Chapter V Explanation of Context Evaluation

When to Use It

- To examine the climate and culture of the environment in which NET operates at the State level; and
- To examine the situation in which major projects or programs operate at the school district, school, or sponsoring agency level.

How to Design It

Lends itself to qualitative methods including:

- Direct observation,
- ## Ethnography,
- Tnterviews,
- Focus groups, and
- Cognitive response techniques.

Who is Targeted?

Administrative and management personnel in:

- State agencies,
- School districts,
- Sponsoring agencies, and
- Cooperating organizations.

Introduction

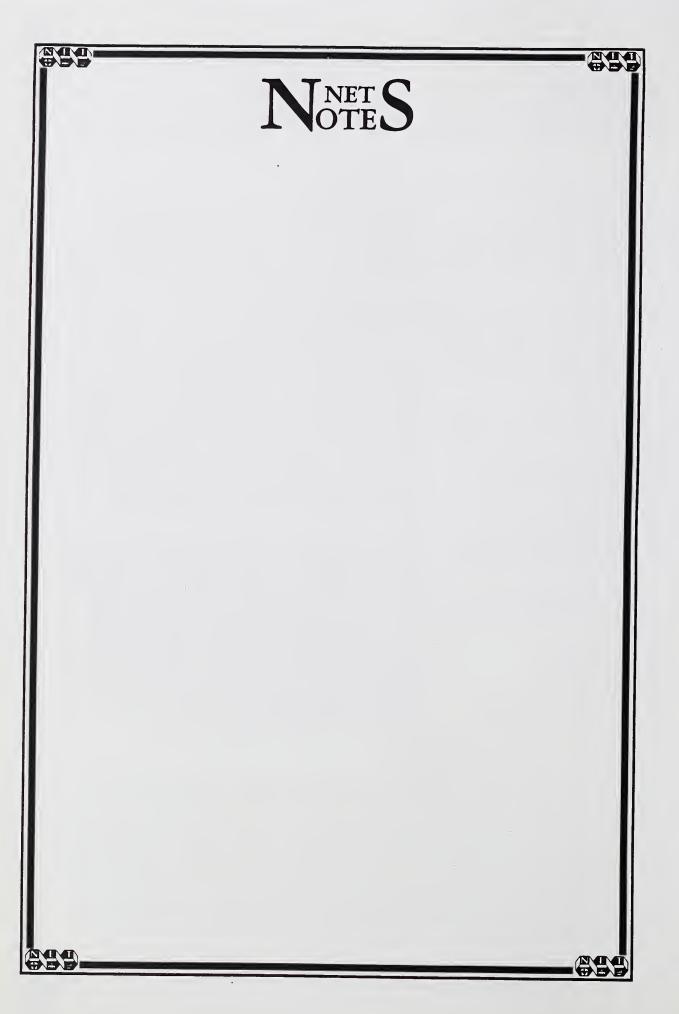
Context evaluation focuses on the overall social, political, economic, and physical environment in which the nutrition education and training program is planned and implemented; the goals and objectives of the program; and how the goals are perceived by different individuals. It answers the primary question, "In what way(s) did the context in which the project or program was conducted affect the achievement of the goals and objectives?" Extensive research has shown that contextual factors such as organizational climate and culture significantly affect implementation of any new approach. Unfortunately, these important considerations are generally ignored in program evaluations, particularly in evaluations of staff development (Guskey and Sparks, 1991).

Context evaluation focuses on the macro-environment of the program, which, in many ways, establishes its scope of influence and public image. It addresses how external factors shape and mold the purpose and comprehensiveness of the State's Child Nutrition Program, and, in turn, help to shape the perspective of various individuals relative to the importance and purpose of the program. Those perceptions become a major contributing factor to defining the parameters, goals, and objectives of a given program.

Why Address this Type of Evaluation?

Context evaluation helps a NET Coordinator understand the situation (context) in which the Program is to bring about change. One important area of study in context evaluation may be determining how State, district, and school administrators view nutrition education and training, and the degree to which they display support and commitment for the program. If, for example, extensive staff development programs have been conducted by the State agency to develop collaborative planning efforts among faculty and school nutrition personnel, but the school principals fail to provide joint planning time in the school schedule, the probability of the staff development program enhancing collaborative planning efforts is restricted regardless of how good the







program might be. This is an example of how a State training program can be adversely affected by the context in which implementation occurs at the local level.

In this example, if the evaluation only focused on the resources (input), the implementation process, or participant outcomes of the staff development program, the root cause of the problem may go undetected. If, on the other hand, this type of evaluation data had been provided, the program planners could, in the next planning cycle, identify and offer strategies to address the school administrators' perspective prior to offering the staff development program.

When to Use Context Evaluation

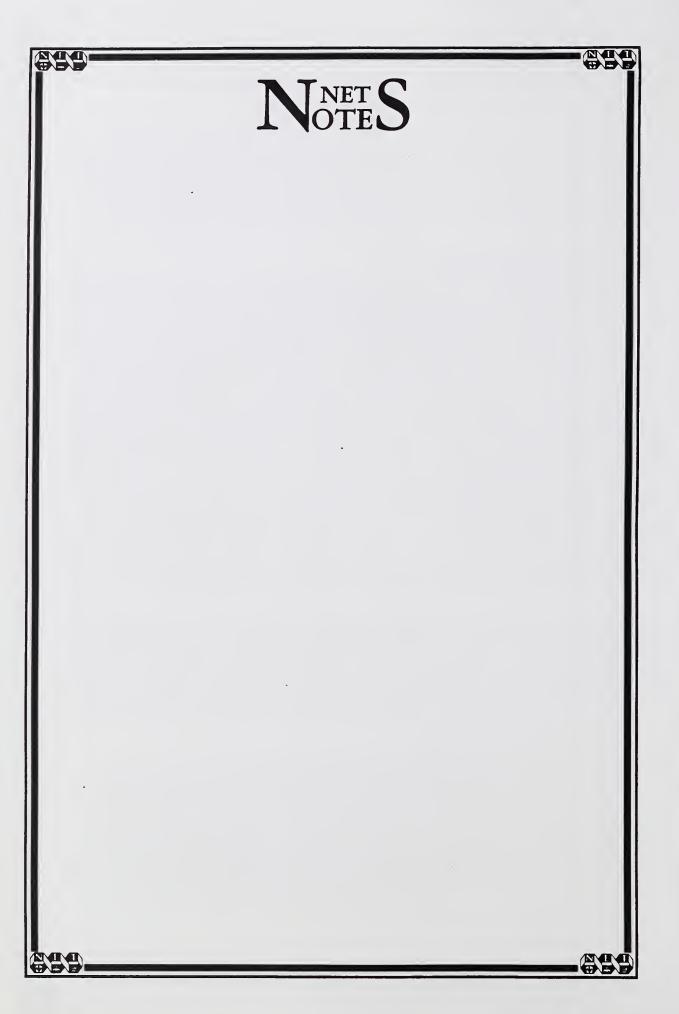
Context evaluation provides the NET Coordinator information about the climate and culture of the environment in which a program operates. This is important in evaluation of the State NET Program as a whole and it is important in evaluation of major developmental projects. Context evaluation for a training program which focuses on the environment of an organization such as a school, a school district, or a State training system, provides decision makers with information regarding how that environment influenced the success or failure of the training program. For example, the success of a State staff development (training) delivery system may be partly attributable to appropriate planning that provided for a support network of school nutrition administrators and school nutrition trainers throughout the State. This State network for staff development can provide for an on-going delivery system for all training programs with continuing support to the trainers and participants from the State and school district levels. When the focus of context evaluation is on the overall NET Program, an example of a contextual evaluation question which may be addressed is, "What is the relationship of the organizational location of the NET Program within State government (i.e., as part of the Child Nutrition Program staff, as part of the curriculum staff, or as part of the Department of Public Health) to the perceived importance and the administrative support of the Program by State-level administrators?"

Thus, it is important that context evaluation be addressed in a comprehensive manner since it can provide strategic clues for projecting the ultimate success of a program or examining why a program or project was or was not successful. When contextual questions are addressed in the initial stages of the program design, they may help to prevent failure. If, on the other hand, these types of data are collected summatively, the findings can provide direction for future program planning.

How to Design Context Evaluation

Context evaluation is not an exact science and lends itself more appropriately to qualitative research. The focus can be on the subtleties and nuances contributing to the public image of the program as well as on the complexities which fashion the scope of a program. Direct observation, ethnography, interviews, focus groups, and cognitive response techniques are commonly used methods for collecting data using a qualitative evaluation design. Guidelines and techniques for using observation, interviewing, and focus group methods are described in detail in the Appendix of the Needs Assessment Guide For The Nutrition Education And Training Program (1994, pp. 71-88).







Ethnography is a type of observation methodology that focuses on observing the participants in the context of their daily lives over an extended period of time. Kolasa (1981) suggests that ethnography has all the advantages of direct observation, in addition to producing a "deeper, richer understanding of context and meaning." Although this type of evaluation can be expensive when observers are paid evaluators, it could be the most appropriate method for certain kinds of nutrition education and training projects.

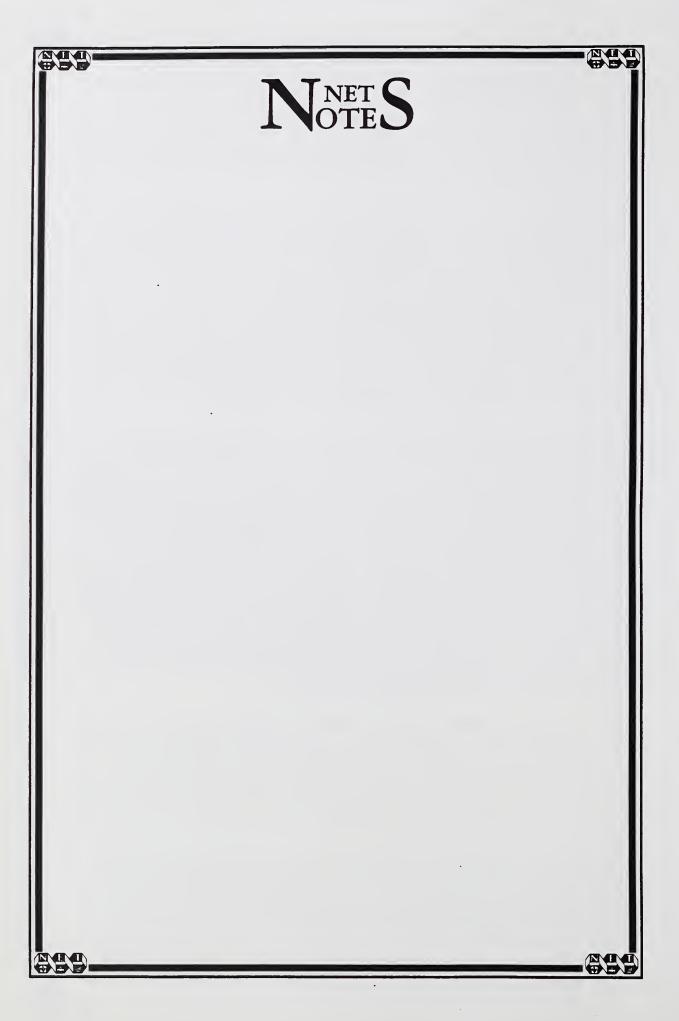
For example, ethnography could be used as an evaluation method in a pilot-project school receiving NET funds to develop a model for school nutrition personnel to work effectively with special education teachers and parents to meet the special dietary needs of those students with diet prescriptions. For this kind of pilot project, where one of the purposes would be to replicate it in other schools, it would be important to describe the context in which the teamwork developed. A trained observer could be funded to observe the interactions of teachers, school nutrition personnel, and parents to identify those behaviors that promoted teamwork, problem solving, and an atmosphere of support and trust. The usefulness of context evaluation in this example is obvious. The project could not be replicated in another school, no matter how good the model, without having or developing the same kind of organizational climate in the school.

Another method appropriate for context evaluation is the cognitive response method. This method requires respondents to "verbalize the thoughts, feelings, and ideas that come to mind while examining or reading a message" (Achterberg, 1988). This method can be especially helpful in trying to determine the attitudes of different groups relative to the importance of nutrition education and training and/or to the goals and objectives of particular programs. For example, this method might be appropriate for use in obtaining context evaluation data from school administrators regarding the issue of allowing the sale of snack foods in the school and where the proceeds should be used. A series of pictures could be shown to a sample of principals to obtain their reaction. Pictures could include soft drink machines; vending machines with mostly highfat, salty snack foods; vending machines with mostly fresh fruit, reduced fat, and reduced salt snacks; a snack bar run by the athletic department; and sale of candy by school clubs. School administrators' responses to these pictures could provide the NET Program important contextual evaluation data for use in planning nutrition education projects.

In the qualitative methods described, randomness in sample selection is not a concern, and may, as Achterberg suggests, be counterproductive to what is needed (Ibid., 1988). Sampling for these qualitative evaluation methods is focused on inclusiveness and on selecting individuals who exhibit a full range of characteristics of the population. Since standard statistical tests are rarely used, sample size is not as important as the inclusiveness of characteristics, thus, sample size may be limited to a relatively small group of 10 to 20 people.

Once the data have been collected, they are aggregated into broad, topical categories. Next, a theme analysis is conducted by studying the data to identify the presence of any underlying themes. These analyses are then used to construct a theoretical framework which attempts to synthesize the data into a meaningful whole. This produces what is referred to as a "grounding theoretical framework which attempts to synthesize the data into a meaningful whole.







ry" which helps to ensure the validity of the data and provides a framework for describing the findings (Glaser and Strauss, 1967). Consider the previous example of the cognitive response method used to obtain contextual evaluation data from school administrators regarding their attitude toward the sale of snack foods in their school. Data gathered from principals throughout the State would be aggregated into broad categories such as: primary concern with nutritional integrity in the school, primary concern with meeting student expectations, or primary concern with obtaining extra funds to pay for school needs. The three underlying themes then could be described as, "nutrition, student requests, or money for school." This type of context evaluation would be useful to State planners as well as school district nutrition directors for designing training programs for school administrators and for preparing to meet the challenge of ensuring nutrition integrity of school food service programs.

Another type of qualitative research design is the data review method. If the evaluator is interested in determining the demographic profile of the State or of a given region, with a particular focus on the population's societal and political orientation, the review of current demographic studies and statistics may prove quite valuable. Information on this design can be found in the Appendix of the NET Needs Assessment Guide (1994).

Summary

Context evaluation involves the examination of the situation in which a program or project operates. It lends itself to qualitative evaluation methods and generally targets program administrators in State agencies, school districts, sponsoring agencies, and cooperating organizations.

Note S 829





Chapter VI Explanation of Input Evaluation

When to Use It

- As part of needs assessment to examine resources available and competency of instructional staff;
- During program implementation (formative), to examine adequacy of resources, entry level ability of students, and competency level of instructors; and
- Following program implementation (summative), to examine adequacy of resources, entry level ability of students, and competency level of instructors.

How to Design It

Qualitative methods include:

- Direct observation,
- Ethnography,
- Interviews,
- Focus groups, and
- Expert judgment.

Quantitative methods include:

- Survey questionnaires,
- Structured interviews, and
- Experimental design.

Who is Targeted?

- Program administrators and management personnel;
- Individuals involved in implementation, e.g. trainers in a staff development project; and
- Program participants.

Introduction

Input evaluation focuses on the appropriateness and adequacy of all human, fiscal, and physical resources that go into the design and provision of a given program or initiative. Included are such resources or input as equipment, facilities, materials, supplies, curriculum guides, entry level ability of students, and the competency level of the instructional staff providing the education or training program. In other words, it focuses on everything that goes into the makeup of a given program or project.

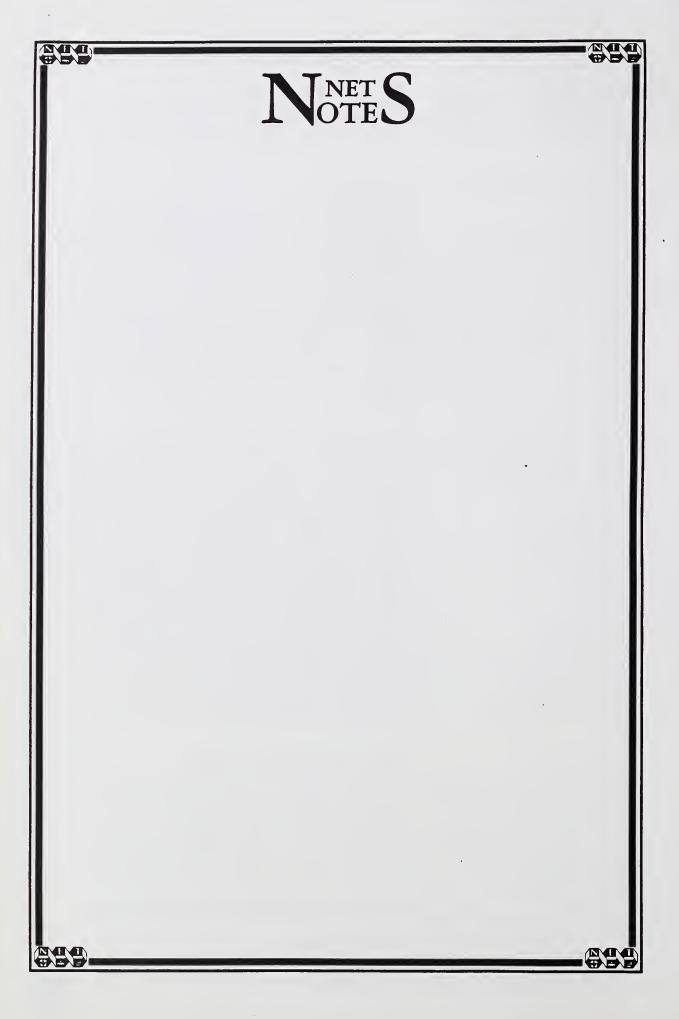
Why Address this Type of Evaluation?

Input evaluation can make a major contribution to decision making by providing information useful for examining alternative strategies, designs, and resource needs for the NET Program or its projects. For example, input evaluation could provide important data regarding the use of a State nutrition education curriculum guide which calls for extensive use of computer software in classroom settings where the instructors lack computer skills, or the classrooms do not have adequate computers for student use. Input evaluation for a State training program provides valuable insight regarding the competency level of the State cadre of trainers. If they do not have the skills necessary to present a training program, the quality of training is compromised at the outset.

When to Use Input Evaluation

Valuable information can be provided through input evaluation, enabling the program planner to design programs and provide resources necessary to ensure the success of the program. Even the brightest students cannot succeed if the program design is faulty or if the appropriate type of equipment is not provided. Input evaluation can help the planner avoid these types of costly mistakes or can provide data for modifications to avoid them in the future.







As part of a needs assessment, input evaluation can be used to help the NET Coordinator determine whether resources are adequate for the job to be done, and whether the most appropriate resources are selected and made available for use in the program. Further, it helps ensure that an appropriate curriculum or instructional design is used, that instructors are selected on the basis of their qualifications to present the specific program, and that students have the appropriate prerequisite abilities to enter and succeed in the program as it is designed.

When input evaluation is conducted during implementation (formative) of a project or program, it can provide feedback on changes that need to be made immediately to help remedy deficiencies and ensure success. When input evaluation is used at the end of a program (summative), it can provide some revealing clues as to why some programs failed while others succeeded. Some examples of input evaluation questions for a school nutrition education program are shown below.

"Did the curriculum guide include the types of instructional activities and design to ensure student success in meeting the objectives of the nutrition education program for grades 4-6?"

"What kinds of instructional materials were most appropriate for use in this program?" "Did the type of equipment and facilities provided help ensure success in meeting the objectives of the program?"

"Did the teachers have the appropriate skills to use the curriculum guide and the equipment and facilities provided?"

"Did the students have the requisite skills when they entered the program to ensure they were able to succeed?"

How to Design Input Evaluation

Input evaluation data may be obtained during implementation (formative) or at the end of a program or initiative (summative). It can be used with a qualitative design using methods such as observation, interview, ethnography, expert judgment, or a quantitative experimental or quasi-experimental evaluation design. Guidelines for using observations and interviews are described in the Appendix of the NET Needs Assessment Guide (1994, pages 71-87).

Expert judgment may be used as an evaluation method appropriate for input evaluation. One approach is an expert review from an individual who is given an open-ended opportunity to assess the probability of a program's success given the input factors assigned to it. Another approach would involve several individuals working independently to make an assessment. Expert judgment may also be obtained from a group of individuals using a structured technique such as the Delphi technique (Cotler, 1988, p. 273).

An experimental design could be used appropriately for input evaluation, however, this type of evaluation entails a more sophisticated approach than most other methods described in this Guide. This type of evaluation design is usually used by researchers in a university setting who have been contracted to develop and manage the evaluation. For example, if the evaluation focus was on the design of a course, a comparison might be made between the use of multi-media in one class and a more traditional course design in another. Using an experimental design incorpo-



NoteS 829



rating a control group and an experimental group, comparison could be made of the differences realized between the groups based on one group's use of multi-media materials (e.g. audio, video, animation graphics) in a training program as compared to the other group's use of traditional printed materials (a training workbook). An experimental design may focus on a training program that incorporates experiential learning activities in every class session as compared to one that incorporates only reading, lectures, and class discussions. For an explanation of experimental design, refer to a textbook on research methods such as *Understanding Research Methods* (1985) or *Experimental Design* (1982).

Summary

Input evaluation provides information on how to best use resources to accomplish program goals and objectives. These data can be obtained during implementation of a program or project (formative) or at the end (summative). Qualitative and quantitative data collection methods can be used appropriately for input evaluation.

Note S

849





Chapter VII Explanation of Process Evaluation

When to Use It

During implementation or after program completion:

- to identify problems in procedural design,
- to identify how well the students are performing or performed on interim performance objectives, and
- to assess to what extent the instructor or trainer is using or used the materials as they were designed.

How to Design It

Qualitative methods include:

- Direct observation, and
- Expert judgment.

Quantitative methods include:

- Survey questionnaires,
- Structured interviews,
- Ex Tests,
- Authentic performance assessment, and
- Experimental design.

Who is Targeted?

- Program administrators and management personnel;
- Individuals involved in implementation, e.g. trainers in a staff development project; and
- Rogram participants.

Introduction

Process evaluation focuses on whether the resources for the program are used appropriately during the implementation phase, if the program is implemented as it was designed, and if the students are meeting interim performance objectives. Whereas input evaluation addresses the appropriateness and adequacy of resources, process evaluation asks whether they are used as they were designed. If the focus is on the evaluation of a training program for staff or a nutrition class for students, process evaluation is most closely aligned with the instructional dimension of the program. If the focus is on resource utilization, however, process evaluation may be designed to focus on the management of the program as well as on the instructional process.

Why Address this Type of Evaluation?

Process evaluation gets at the "heart and fiber" of the program and helps provide answers as to what specific features during program implementation contributed to success or, conversely, why the plans went awry. This type of evaluation describes what actually takes place during the program or project. Therefore, it can provide valuable information for making mid-course corrections in a program or project in a timely manner.

For example, a one-day training program for Child and Adult Care Food Program (CACFP) Sponsors offered in eight locations throughout the State over a year's time should be designed so that process evaluation data are obtained after each one-day program offering. These data can help planners make necessary changes in the program design during implementation in order to improve its effectiveness. Thus, if after the program is presented

twice, the process evaluation indicates that the trainers are not using the training plan as it was designed, the NET Coordinator should meet with the trainers to ensure that changes are made. This may involve some re-training or coaching of the trainers.





One of the major benefits or contributions of process evaluation is the identification of flaws or problems in procedural design, whether the design relates to the instructional process or to the managerial aspect of the program. It helps provide an answer to the question, "If we had it do to over again, what would we do differently?" The answer to this question can provide valuable input in the redesign of a program or management procedure.

Process evaluation lends itself to both quantitative and qualitative evaluation methods. Traditionally, most of the process evaluation studies have utilized quantitative, experimental designs where the primary objective was the measurement of differences. Consider a nutrition education project which focuses on students' acquisition of new knowledge about nutrition and how nutrition affects health. The evaluation design includes pre- and post-test measures for a given unit of study. The process evaluation design for this project might include explicit measures on how frequently specific instructional materials were used, in what manner these materials were used, with what age groups of students certain materials were used, what types of instructional competencies teachers needed to use the materials most effectively, and what differences these factors had on the student's acquisition of new knowledge.

Qualitative designs, on the other hand, might focus more on the feelings, attitudes, general concerns, and issues that surround and/or permeate the process, and could provide meaningful clues as to why the process is or is not working. For example, the focus might be on the importance of multi-cultural sensitivity and its effect on the attitudes of students in nutrition education classes. As our population becomes more diversified in its ethnicity, it is important that ethnic differences, cultures, and customs be recognized and valued. Thus, when the Dietary Guidelines are taught, an evaluation question might ask how teachers explain the Food Guide Pyramid using food examples which illustrate the variety of ethnic backgrounds of their students. This question illustrates the process used in the classroom to present curriculum materials.

When to Use Process Evaluation

Process evaluation always addresses what happens during the implementation phase of a program or project. The data may be gathered *during* the implementation phase or they may be obtained *after* the implementation phase. In either case, the evaluation question asks "how or what" about the process of implementation. Following are some examples of process evaluation questions which could be addressed during implementation or following implementation.

- * "How are the curriculum materials being used in the instructional process?" (formative evaluation)
- "How were the curriculum materials used in the instructional process?" (summative evaluation)
- "How are the equipment and facilities being used in the process of implementing the program?" (formative evaluation)
- "How were the equipment and facilities used in the process of implementing the program?" (summative evaluation)





- Is the instructor utilizing the variety of instructional techniques appropriate for presenting the program?" (formative evaluation)
- "Did the instructor utilize the variety of instructional techniques appropriate for presenting the program?" (summative evaluation)
- "How well are the students performing? Are they meeting interim performance objectives?" (formative evaluation)
- "How well did the students perform on the interim performance objectives?" (summative evaluation)

The focus of process evaluation might also center on other aspects of the overall NET Program. It might, for example, focus on the management of the program itself. From this perspective, process evaluation for the NET Program might address such questions as:

- "How well are we staying within budget?"
- "How well are we following appropriate procedures in managing external contracts to ensure delivery of contracted products and services?"

How to Design Process Evaluation

Since process evaluation assesses what takes place during implementation of a project or program, a variety of forms of feedback can be used. The methods used most often include written questionnaires, tests, interviews, and observations. Assessment of interim student performance, whether for a K-12 classroom or for an adult learning situation, is an important type of process evaluation.

Traditionally, we have:

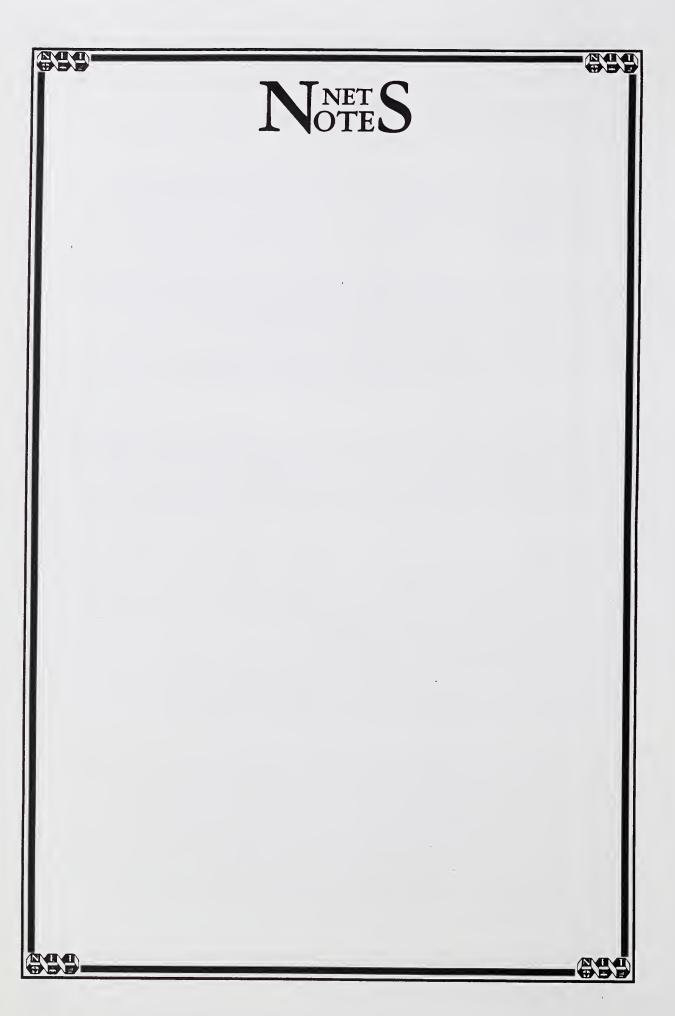
- Example 2 Compared learners to each other and placed more importance on those comparisons than on what was actually learned,
- See Used paper and pencil tests and checklists to evaluate what students learned, and
- Excluded students from the evaluation process.

Contemporary assessment of performance requires us to:

- 🕸 Have the students take an active role in evaluating their own progress,
- 🕸 Use a variety of sources and tools to determine what students have learned, and
- 🕸 Evaluate students in relation to authentic criteria of actual performance.

There are several types of process evaluation used in student classrooms and adult training situations. For K-12 students, the move today is away from paper-and-pencil tests for either process or product evaluation, and toward the use of evaluations where the learner can demonstrate use of the new knowledge and skills. This move to a more authentic type of assessment is also true for adult training programs. Performance assessment is one of the most useful types of evaluation in student classrooms as well as for job-related training. This type of assessment requires the learner to actually demonstrate the skill or use of the knowledge. For K-12 students as well as for adult learners, this type of assessment is more appropriate than paper-and-pencil tests because it requires the learner to demonstrate learning in a situation similar to the one in which it will actually be used.







Authentic performance assessments may take many forms, such as:

- models.
- videos,
- debates,
- exhibits,
- diagrams,
- research projects,
- teacher observation,
- · anecdotal records,
- community projects,
- structured activities/projects,
- demonstrations,
- displays,
- reports,
- journals,
- simulations,
- projects,
- presentations,
- case problems, and
- concrete representations.

The move today, for student classroom instruction and adult learning situations, is to authentic performance assessment. This type of assessment focuses on the identification of evaluative criteria and strategies to determine whether or not a student can actually demonstrate his/her proficiency to perform a given objective in a real-life context. Standardized tests or the typical truefalse, matching, completion, or multiple-choice tests usually assess student learning at the comprehension level of cognition, often requiring only recall of facts. Authentic assessment, however, focuses on the use of criterion-referenced tests or evaluation measures which are designed to assess learning at a higher level, from the application up to the evaluation level. Criterion-referenced measures require absolute congruence with the performance objective and objectivity so as to ensure a common understanding between the student and the instructor as to what is being actually required and the measures by which the performance will be assessed.

Use of alternative assessments in schools represents a fundamental shift from reliance on standardized testing techniques that have characterized student assessment for decades (Wolf, Bixby, Glen, & Gardner, 1991). Although there are various alternative assessment techniques, they are generally characterized by:

- The student developing products rather than recognizing answers,
- The student demonstrating a depth of knowledge rather than breadth, and
- The teacher using his/her expertise to judge the assessment rather than using a specific mechanical standard (Calfee and Perfumo, 1993).

These characteristics could be used as guidelines for developing authentic assessment measures for both K-12 students and adult learners.



Note S



One of the most popular forms of alternative assessment is the use of portfolios. Portfolio assessment is appropriate to use for process evaluation and for product evaluation. Refer to Appendix D for a description of portfolio assessment.

Summary

Process evaluation provides feedback about program implementation. It is usually conducted during implementation, but may also examine the process of implementation after the program or project has been completed. Assessment of interim student performance is an important aspect of process evaluation. One of the most effective approaches to use for this type of evaluation is authentic performance assessment.





Chapter VIII Explanation of Product Evaluation

When to Use It

Following program implementation to describe students performance in relation to stated performance objectives (measurable outcomes).

How to Design It

- ✓ Authentic performance assessment,
- Tests.
- ZRating scales,
- Attitude questionnaires,
- **Z**Observation,
- Case study, and
- Experimental design.

Who is Targeted?

- Adult training program participants, and
- ZK-12 students.

Introduction

Product evaluation focuses on student outcomes, ability levels, and performance after the completion of a given program. This type of evaluation may be referred to as outcome evaluation. Instructional designers sometimes define performance objectives as interim or enabling objectives, which would be evaluated in process evaluation, and terminal objectives, which would be evaluated in product evaluation. Regardless the name it is given, this type of evaluation addresses what students know and are able to do upon completion of the program. Perhaps, more importantly, it focuses on the degree to which students are able to apply the new knowledge in real life settings. Likewise, product evaluation focuses on what staff personnel know and are able to do upon the completion of a staff development training program. Thus, product evaluation seeks to determine the demonstration of behavioral change.

Why Address this Type of Evaluation?

Unless product evaluation is viewed as one of the most essential components within a multi-faceted evaluation

design, no real relevancy or meaning can be given to context, input, and process evaluation. When context, input, and process evaluation are determined at the end of a program or project (summatively), the data are examined in relation to student outcomes and/or the outcomes of the program. Some examples are shown below.

A context evaluation question: "Did this type of administrative support or environment contribute to the success of the program?"

An input evaluation question: "Did these types of resources and the instructional design contribute to the success of the program?"

A process evaluation question: "Did the implementation of the program in this manner contribute to the success of the program?"

Thus, product evaluation provides critical information for judging the success of a program or project. Knowing that the administrative staff were supportive, that the program was needed, that all of the requisite inputs were provided, and that the program was implemented as planned still does not address the question: "But can the students/staff actually do these things?" That question can only be answered through a carefully designed product evaluation.



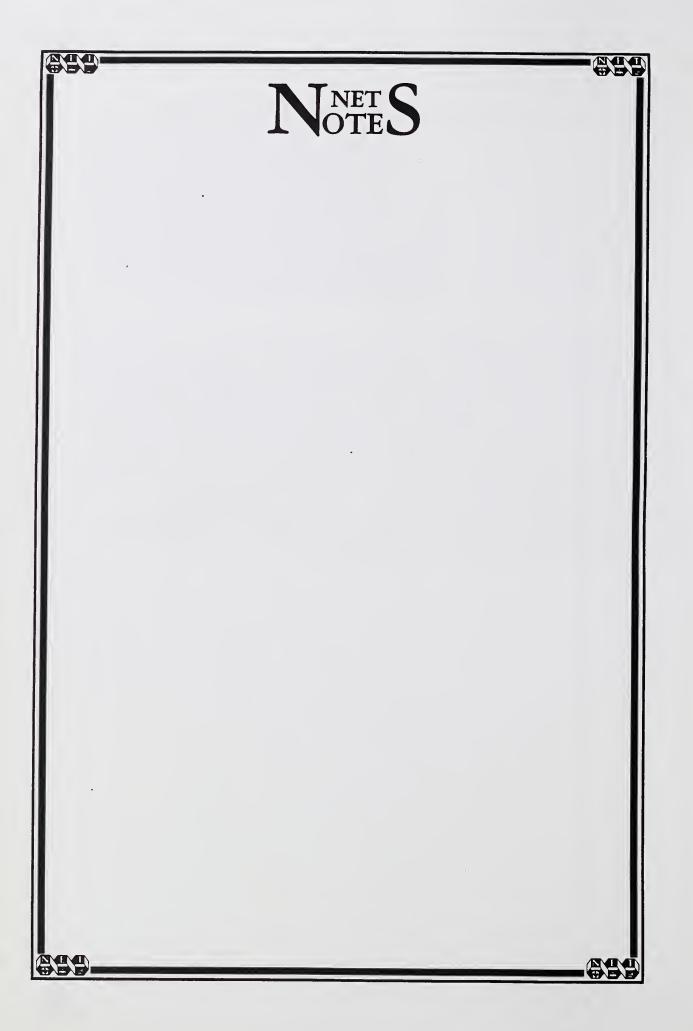
When to Use Product Evaluation

Just as evaluation closes the loop in the Assessment-Planning-Budgeting-Implementation-Evaluation Process, product evaluation closes the loop in the curriculum and instructional design process. The curriculum planner or instructional designer begins with an assessment of what students need to know and are demonstrating as their typical behavior or performance in comparison to what they should know and demonstrate in their daily lives or job performance. Based on that assessment, the design phase of the instructional process is begun. Once designed, the curriculum is implemented through the instructional process, and the ultimate outcomes of that process are assessed. The ultimate performance outcomes are assessed as product evaluation. Based on the evaluation of the students' performance upon completion of the program, the needs assessment begins again and the curriculum is re-designed or modified for improvement.

How to Design Product Evaluation

Product evaluation focuses on what the students can actually do. This raises two critical questions: "What level of learning or cognition is required?" and "When, or at what point in time after program completion, will student performance or behavioral change be evaluated?" The answer to the first question requires a familiarity with the levels of learning in the cognitive domain, as well as familiarity with alternative application designs. The study of cognition is not an exact science because, as yet, we do not truly understand how the brain and the mind work. However, most learning theorists and curriculum planners recognize the existence of a learning continuum or taxonomy, within which several levels of learning have been generally described.

To illustrate the learning continuum or taxonomy, consider an example of a student learning experience that is focused on the Food Guide Pyramid (See Illustration 3. An example of the learning continuum with a focus on the Food Guide Pyramid.).



Evaluation.
The student
may be asked
to make several
recommendations
for improving the
weekly menu so as to
assure the inclusion of the
recommended number of
servings every day and to
assure variety in each food group
every day.

Synthesis. The student may be given a weekly menu from the cafeteria to determine the composite picture for the week, including the identification of the variety of foods used to meet the recommended number of servings on the various levels of the Food Guide Pyramid.

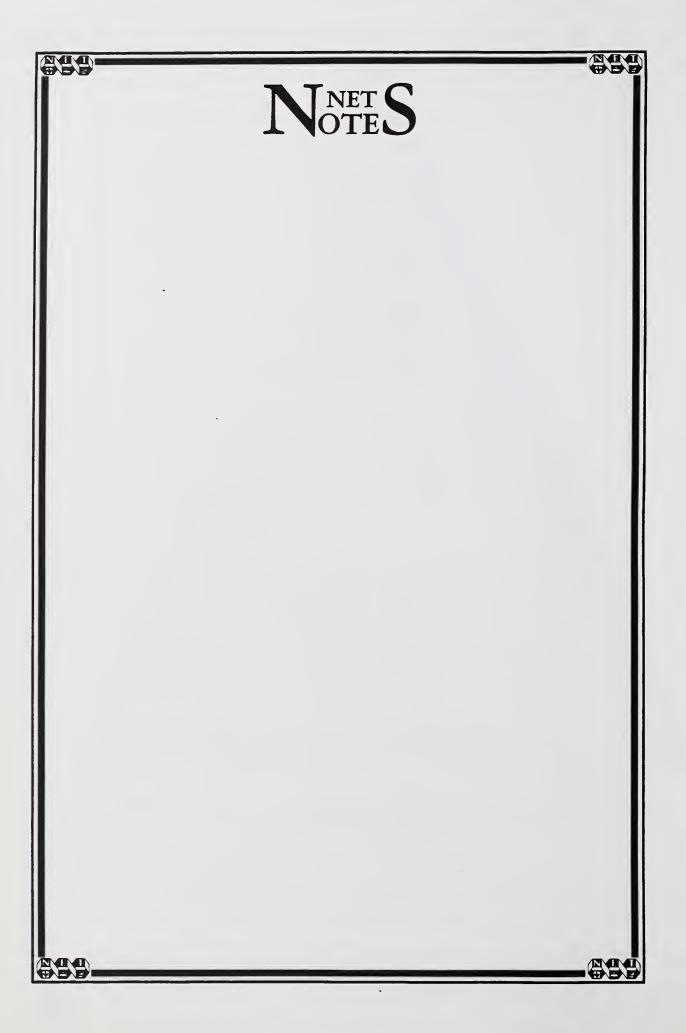
Analysis. When given a sample daily menu, the student would be able to determine whether the daily menu is consistent with the recommended number of servings on the various levels of the Food Guide Pyramid.

Application. The student is able to use the information or knowledge in some limited way, e.g. identify what he/she might eat for one day to include at least one selection from each level of the Food Guide Pyramid.

Comprehension. The student may recall definitions and recognize examples of or differences between categories or groups, e.g. match examples of different foods from a given list to the appropriate food groups in the Food Guide Pyramid.

Knowledge is the lowest level of cognition. The student is only able to recall facts. e.g. name the groups of foods in the Food Guide Pyramid.

Illustration 3. An example of the learning continuum with a focus on the Food Guide Pyramid.



This example shows that the knowledge level of the student and the ability to apply the knowledge increases at each level or phase along the continuum. Thus, if the student's performance level is expected to change, the curriculum and instructional design *must* move the student at least to the application level or higher. Curriculum designs and tests which focus solely on the recall of facts and on such outcomes as naming, listing, describing, defining, and matching, will never effect behavioral change. Meaningful product evaluation, therefore, must focus on what the students can do at the application level of cognition or higher.

Another aspect of product evaluation is that of application designs, which require the curriculum planner to examine the variety of contexts within which the application of learning should be demonstrated. Again, there are several levels of contexts, each requiring a higher level of proficiency. The example below shows several different contexts, each requiring a higher level of proficiency to be demonstrated by a school nutrition manager in a training course on implementing the Dietary Guidelines.

NET

NET

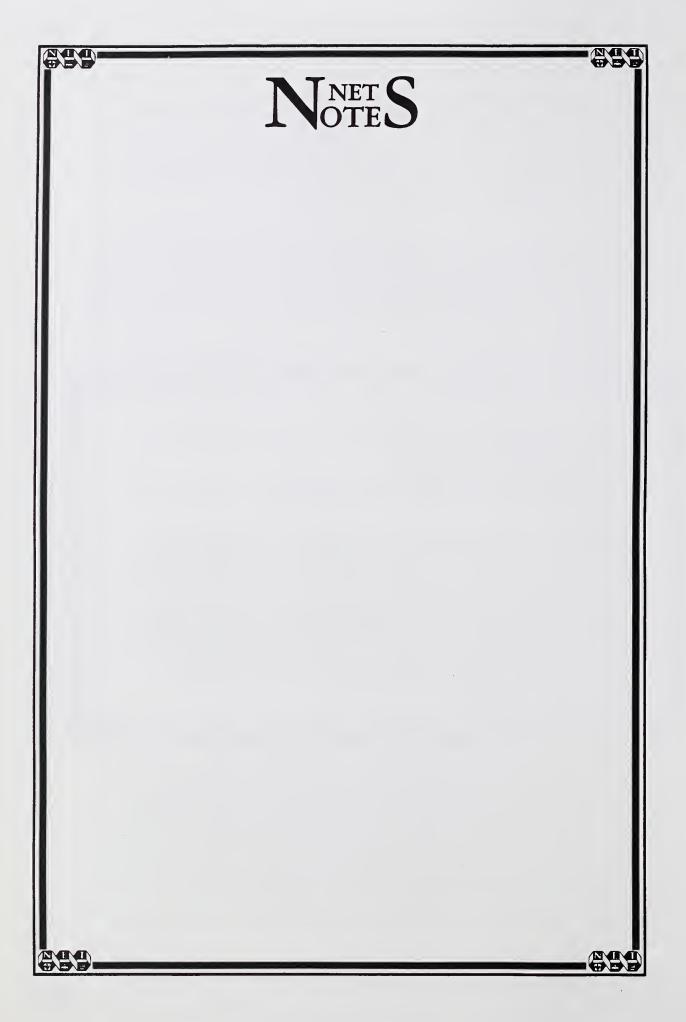
Application is demonstrated in the abstract with little or no relationship to real problems or situations, e.g. simple addition of the fat grams in the foods in a meal given as a written math problem.

Application is demonstrated within a meaningful context, e.g. calculation of the total fat content for a real meal that an individual student selected and ate on a specific day.

Application is demonstrated within a meaningful problem-solving context, e.g. after the manager determines whether the example student are a large amount of fat in the selected lunch, the manager is asked to outline changes he/she can make in the menu offerings and ways the food is prepared to resolve that problem. Application is demonstrated within a different, but relevant, problem-solving context, e.g. the manager may find that he/she has a high cholesterol problem (or other type of relevant health problem) and be asked to outline substantive changes to be made in his/her daily eating habits to reduce fat and promote improved nutritional health.



If the instructional designer wants to effect substantive behavioral change on the part of the learners, the design of the curriculum must guide the students to the level of performance actually required in a real life setting. An analogous case or argument can be made in staff development, where the focus is on training instructors or food service personnel. If the objective is to change the demonstrated performance or behavior of the staff, the training program must be designed accordingly. Since product evaluation must be tightly interfaced with instructional design, and the terminal objectives included therein, the product evaluation must focus on what the student or staff member can do in the context in which that behavior or performance is



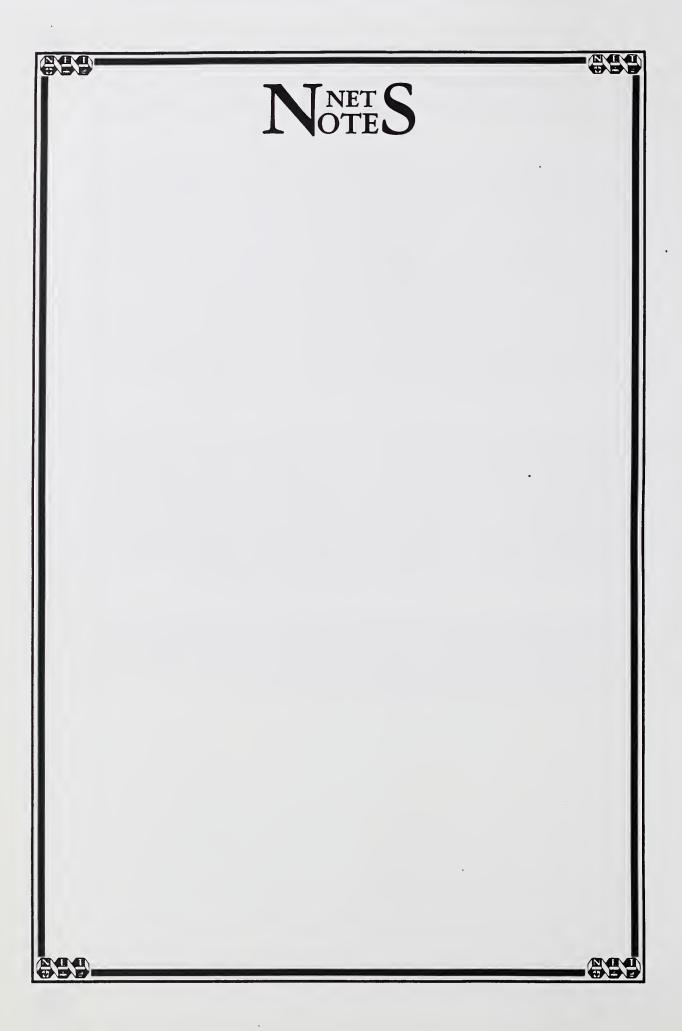
identified in the curriculum objectives. Thus, for the child nutrition professional, the context for performance assessment is the actual job situation. This is the true essence of authentic performance assessment.

The second question the instructional designer must also consider is the time factor in regard to when product evaluation should occur ("When, or at what point in time after program completion, will student performance or behavioral change be evaluated?"). St.Pierre (1982) refers to this as "proximity of outcomes." Should the evaluation be conducted immediately upon completion of the instruction or program, or should the evaluation be conducted at some distant time in the future? In relation to product evaluation, the answer is immediately upon the completion of the instruction or program. This, then, leaves the question of immediacy to be defined. Unfortunately, there is no clearly delineated, objective definition which can be used in every situation. The product evaluation may be on the last day of class, the day after the course or program has been completed, or six months after program completion. Perhaps the definition can most appropriately be determined by answering the question "can they?" "Can the learners demonstrate the level of proficiency required?" And "how does one know that they can?"

When performance objectives are written in measurable, understandable terms, the instructional designer has essentially defined what performance must be demonstrated by the learners. The main decision left is to determine the most effective way to measure the level of performance of the learner. Product evaluation determines the level of student proficiency on defined performance objectives and can be designed using the same methods of evaluation as used for process evaluation. Authentic performance assessment, one of the most useful types of evaluation has already been described in Chapter 7, Explanation of Process Evaluation. Portfolio assessment, one of the most popular forms of authentic assessment, is described in Appendix D.

Summary

Product evaluation measures the attainment of performance objectives (measurable outcomes). When objectives are written in measurable, understandable terms, the product evaluation has essentially been defined. One of the most useful methods for product evaluation in the classroom or adult staff development program is authentic performance assessment.







Chapter IX Explanation of Impact Evaluation

When to Use It

As an examination of the long-term impact of the program.

How to Design It

To evaluate the impact of training programs on behavior of participants:

- Interviews,
- Focus groups,
- Self-assessment questionnaires,
- *Observations,
- *Case studies, and
- *Portfolio assessment.

To evaluate the impact of training programs on the organization:

- Interviews,
- Focus groups,
- *Analysis of documents,
- *Observations,
- *Case studies, and
- *Cost-benefit studies.

To evaluate the impact of nutrition education programs on students' food-related behavior at school and/or knowledge and attitudes about food:

- *Observation,
- *Attitude survey,
- Criterion-referenced test of food and nutrition knowledge, and
- *Analysis of records or documents.

Who is Targeted?

- Program administrators and management staff,
- Adult program participants,
- Participants' supervisors,
- Students,
- Parents,
- School faculty and administrators,
- Members of the community, and
- Individuals from cooperating agencies.

Introduction

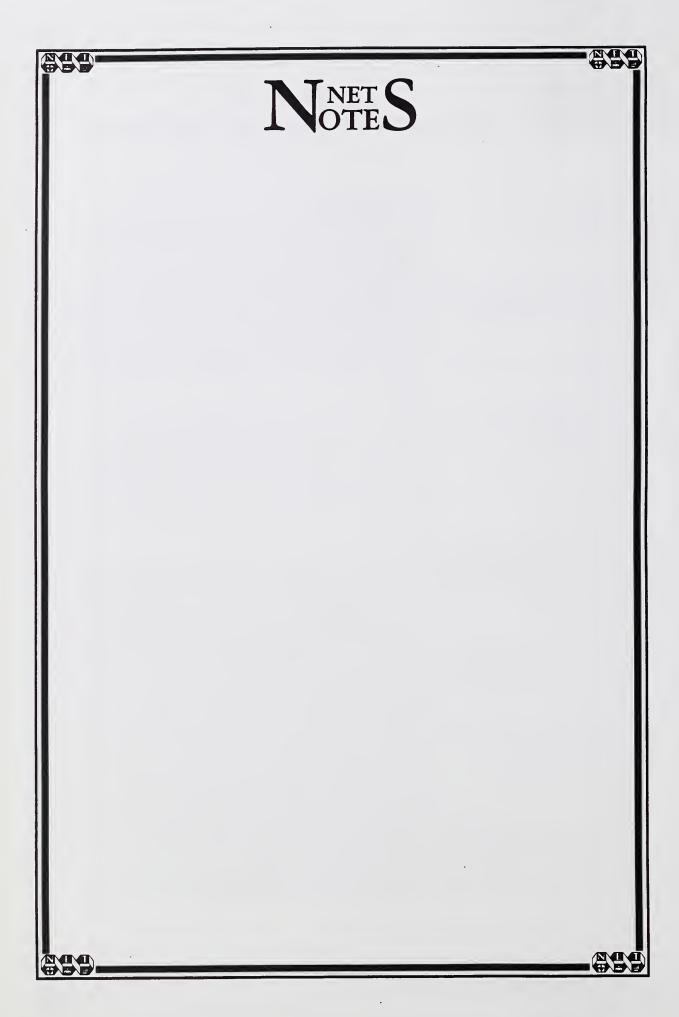
Impact evaluation focuses on the cumulative difference the NET Program makes on the target groups and on the issues or problems it is intended to address. Its focus, therefore, is on long-term impact. Although this type of evaluation is the most complex in design and, to some, the most difficult to conduct, it often provides the basis for critical policy decisions which can affect the continued life and substance of a program.

Why Address this Type Of Evaluation?

This type of evaluation, more than any other, can provide the justification, or lack thereof, for critical policy decisions, including the continuation of a program. "Proof of value," is the essence of impact evaluation.

Some States and some national administrations have, over the years, used a "zero-based budgeting" strategy as a basis for their decision-making process. This strategy requires program managers or directors to justify why their programs should be continued prior to submitting a budget request. The primary and most legitimate focus of that justification must be program impact or proof of value. When administrations change, as they inevitably and continually do, most will begin with a search for ways to cut the current budget in order to find the resources needed to support new programs which they have pledged to initiate. Again, evidence of value, evidence of the difference the program has made and is making, and proof of the program's impact on the alleviation of specific societal problems assumes critical importance. These data can only be provided through impact evaluation.







When to Use Impact Evaluation

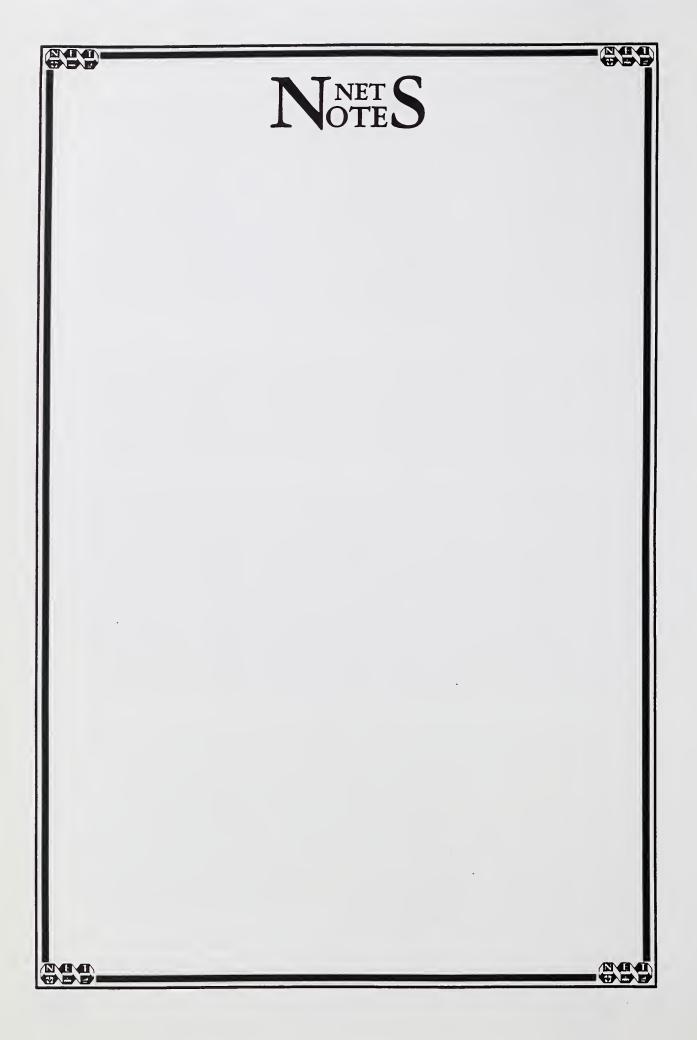
The scope of impact evaluation can be quite extensive, ranging from a behavior change in adult participants in staff development programs, to changes in the Child Nutrition Programs, to changes in the students' food-related behavior at school. Since real change takes time, a NET Coordinator will not be concerned with determining the impact of a two-week unit of study incorporated into the eighth grade curriculum of a school, nor the impact of a one-day or even a several-day staff development session. Rather, impact evaluation takes a more holistic look at the overall effect of an entire program of activity, e.g., the cumulative impact of all NET Program activities within a particular school district, or region, or State over the course of an extended period of time. Impact evaluation, thus, speaks to the other side of the "proximity" issue which was raised previously in product evaluation. The time-line for the evaluation is extended from "immediately" after program completion to a more "distant future."

How to Design Impact Evaluation

Impact evaluation, like context evaluation, lends itself to qualitative evaluation design as well as to quantitative design. Because of the wide array of variables affecting both the type and extensiveness of the identified impact, an empirical approach which seeks to prove direct cause and effect scenarios may be difficult, if not impossible. Thus, attention should also be given to determining "indications of impact" for which a qualitative evaluation design can be quite helpful.

For example, consider an evaluation study designed to determine the impact of a variety of class-room and cafeteria NET activities on students' knowledge, attitudes, and food practices in regard to using the principles of the Dietary Guidelines. Proving direct cause and effect would be virtually impossible, as well as inappropriate, because of the wide array of uncontrolled variables that affect students' knowledge, attitudes, and food practices. Uncontrolled variables could include advertising, parental influence, access to food, time to eat, and many others. However, indications of impact could be obtained from the results of student focus groups before the nutrition education classroom and cafeteria activities on the Dietary Guidelines, and then some months after the activities have been completed. In this example, the impact evaluation would not isolate specific classroom or cafeteria activities to determine what difference they had made. Rather, the impact evaluation would examine the food-related behavior of students at school in a more holistic way.

Because of the comprehensive scope and extended time-frame characteristic of impact evaluation, the evaluation plan should identify an array of evaluation questions which could be examined over a period of time. Changes in the knowledge, skills, and attitudes of adult participants of staff development programs can be measured by evaluation methods such as interviews, focus groups, self-assessment questionnaires, observations, case studies, and portfolio assessments. Changes in the organization, whether a school nutrition program, a center, or a feeding site, can be measured through interviews, questionnaires, analysis of documents, observations, case studies, and cost-benefit studies. Quantitative changes in students' food-related behavior at school can be measured through observation, a criterion-referenced test of food and nutrition knowledge, and an analysis of records such as the food production record.





Summary

Impact evaluation is "proof of value" evaluation. It addresses the cumulative differences the program has made. The focus is on long-term impact. Like context evaluation, impact evaluation questions can usually be answered by collecting data using qualitative methods.

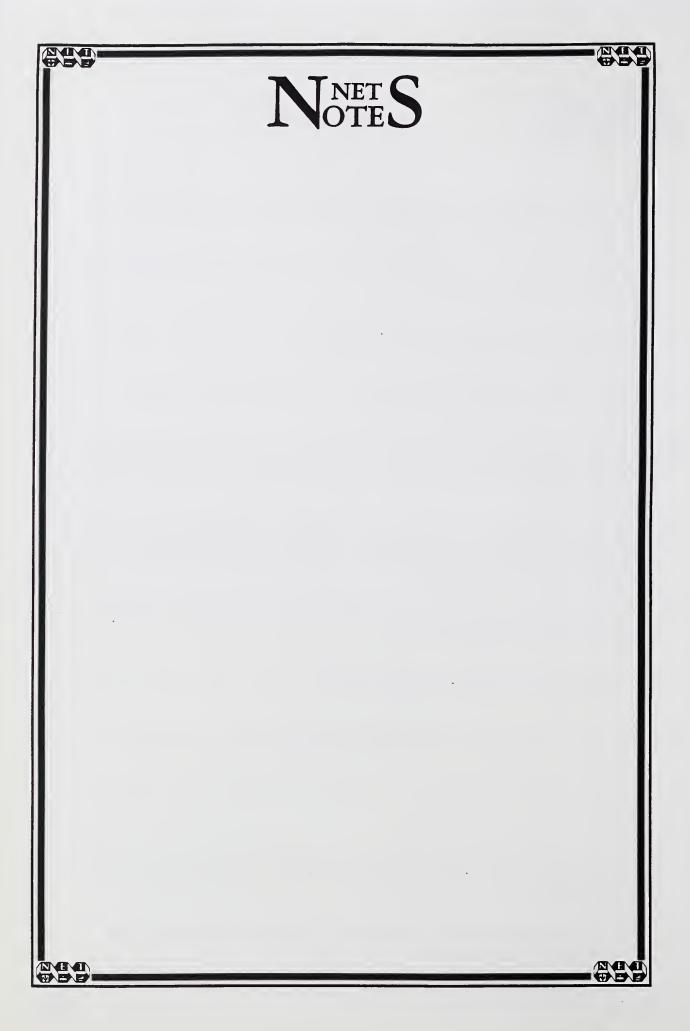






References

- Achterberg, Cheryl. (1988). Qualitative methods in nutrition education evaluation research. Journal of Nutrition Education, 20(5), 244-249.
- American Society for Training and Development. (1991). Statistics for HRD practice, *INFO-Line*, Issue 101. Alexandria, VA: Author.
- Calfee, Robert C. and Perfumo, Pam. (1993). Student portfolios: Opportunities for a revolution in assessment. *Journal of Reading*, 36(7), 532-537.
- Cotler, Phillip. (1988). Marketing management analysis, planning, implementation, and control. Englewood Cliffs, N.J.: Prentice Hall. p. 273.
- Edwards, Patricia K., Mullis, Rebecca M., Clarke, Barbara. (1986). A comprehensive model for evaluating innovative nutrition education programs. *Journal of Nutrition Education*, 18(1), 10-14.
- Filstead, W. J. (1979). Qualitative methods, a needed perspective in evaluation research. In T. D. Cook and C. S. Reichartdt (Eds.). *Qualitative and quantitative methods in evaluation research*. Beverly Hills, CA: Sage Publishing.
- Glaser, B. G. and Straus, A. L. (1967). The discovery of grounding theory: Strategies for qualitative research. New York: Aldine Publishing.
- Guskey, Thomas R. and Sparks, Dennis. (1991, November). What to consider when evaluating staff development. *Educational Leadership*, 73-75.
- Kolasa, K. M. (1981). Nutritional anthropologists and nutrition educators Potential for a multidimensional world view. *Journal of Nutrition Education* 13 (Supp. 1), 9-11.
- Lincoln, Y. S. and Guba, E. G. (1985). *Naturalistic inquiry*. Beverly Hills, CA: Sage Publishing. p. 338.
- Mandell, Jane. (Ed.). (1993). The strategic plan for nutrition education: Promoting healthy eating habits for our children. Alexandria, VA: United States Department of Agriculture, Food and Nutrition Service.
- Martin, Merle P. and Trumbly, James E. (1987, March). A project accountability chart (PAC). Journal of Systems Management, 6-9.





- Mississippi Department of Education, Office of Child Nutrition. (1995). DECIDE to succeed: An orientation for new managers. Jackson, MS: Author.
- Phillips, Jack J. (1991). Handbook of training evaluation and measurement methods. 2nd Ed. Houston: Gulf Publishing Company.
- Robinson, Ann E. and Trapnell, Gail. (1994). Needs assessment guide for the Nutrition Education and Training Program. Alexandria, VA: United States Department of Agriculture, Food and Nutrition Service.
- St. Pierre, Robert G. (1982). Specifying outcomes in nutrition education evaluation. *Journal of Nutrition Education*, 14(2), 61-66.
- St. Pierre, Robert G. and Rezmovic, Victor. (1982). An overview of the national nutrition education and training program evaluation. *Journal of Nutrition Education*, 14(2), 61-66.
- Steele, Sara M. (1991, October). Eight key mental skills in evaluation. Adult Learning, 11-26.
- Stufflebeam, Daniel L. (1971). The relevance of the CIPP Evaluation Model for educational accountability. *Journal of Research and Development in Education*, 5(1), 19-25.
- Trapnell, Gail. (1984, May). Putting the evaluation puzzle together. Training and Development Journal, 90-93.
- Valencia, S. W., with Au, K.H, Schedue, J. A., and Kawakami, A.J. (1990). Assessment of student's ownership of literacy. *The Reading Teacher*, 44, 154-156.
- Wholey, Joseph S. (1991, Summer). Using evaluation to improve program performance. The Bureaucrat, The Journal for Public Managers, 55-59.
- Wolf, D., Bixby., Glenn, J. III., and Gardner, H. (1991). To use their minds well: Investigating new forms of student assessment. In G. Grant (Ed.). *Review of research in education*. Vol. 17, Washington, DC: American Educational Research Association.

For More Information on Evaluation

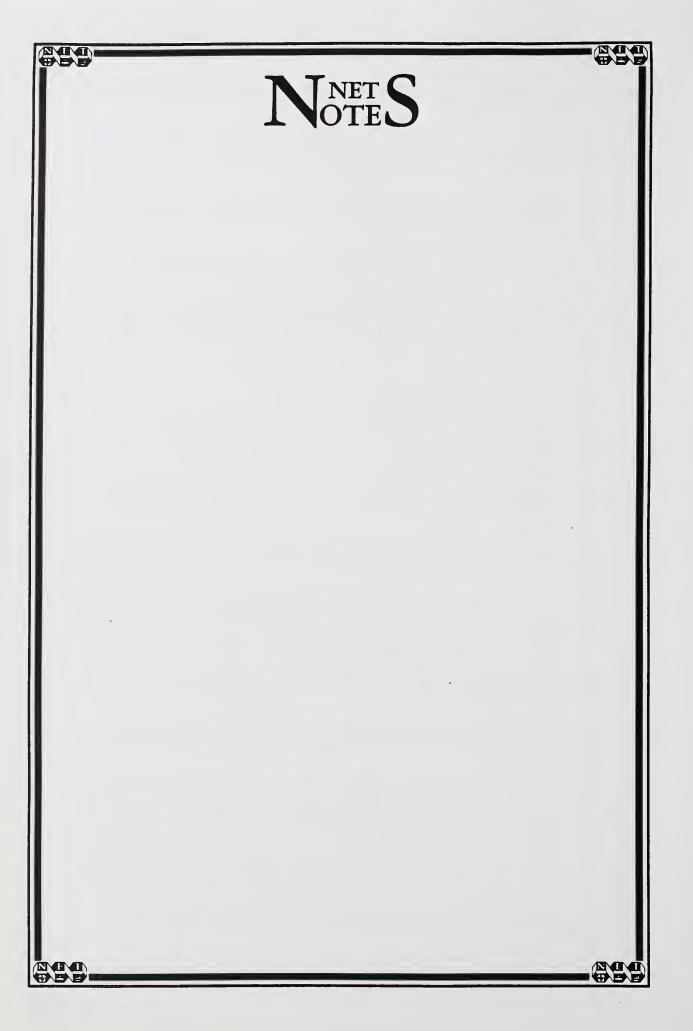
- Adams, Gerald R. and Schvaneveldt, Jay D. (1985). *Understanding research methods*. Longman Publishers: New York.
- American Society for Training and Development. (1991). Measuring attitudinal and behavioral change, *INFO-Line*, Issue 9110. Alexandria, VA: Author.







- American Society for Training and Development. (1990). How to collect data. *INFO-Line*, Issue 008. Alexandria, VA: Author.
- American Society for Training and Development. (1990). How to conduct a cost-benefit analysis, INFO-Line, Issue 007. Alexandria, VA: Author.
- Bakken, David and Bernstein, A. L. (1982, August). A systematic approach to evaluation. Training and Development Journal, , 44-51.
- Ballard, Leslie. (1992, February). Portfolios and self-assessment. English Journal, 46-48.
- Banta, T. W., Cunningham, J. L., Jozwiak, W. W., McCabe, M. P., and Skinner, J. D. (1985). Adapting a nutrition education evaluation to field and political realities. *Journal of Nutrition Education*, 17(2), 41-43.
- Bowman, Barbara. (1987, May). Sins of omission. Training, 45-48.
- Brambley, Peter. (1991). Evaluation training effectiveness: Translating theory into practice. London: McGraw-Hill Book Company (UK).
- Brandt, Ronald. (1978, January). On evaluation: An interview with Daniel L. Stufflebeam. Educational Leadership, 249-253.
- Coffman, Linn. (1990, June). Involving managers in training evaluation. Training and Development Journal, 77-80.
- Cousins, J. B. and Earl, L. M. (1992). The case for participatory evaluation. *Educational Evaluation* and Policy Analysis, 14(4), 397-418.
- Doak, C. C., Doak, L. G., and Root, J. H. (1985). Teaching patients with low literacy skills. Philadelphia: J. B. Lippincott Co.
- Dunn, Samuel and Thomas, Kenneth. (1985, April). Surpassing the "smile sheet" approach to evaluation. *Training*, 65-71.
- Edwards, P. K., Mullis, R. M., and Clarke, B. (1986). A comprehensive model for evaluating innovative education programs. *Journal of Nutrition Education* 18(1), 10-15.
- Erikson, Paul R. (1990, January). Evaluation training results. Training and Development Journal, 57-59.
- Erkut, Sumru and Fields, J. P. (1987, October). Focus groups to the rescue. *Training and Development Journal*, 74-76.





- Finch, C. R. and Crunkilton, J. R. (1979). Curriculum development in vocational and technical education. Boston: Allyn and Bacon, Inc.
- Fisher, Harold E. (1988, January). Make training accountable: Assess its impact. *Personnel Journal*, 73-77.
- Graves, Karen L., Farthing, MaryAnn C., Turchi, Janet M., and Smith, Sally A. (1989). Evaluation of the North Carolina NET Program. *Journal of Nutrition Education*, 21(1), 11-14.
- Hanson, R. A. and Schutz, R. E. (1981). Rethinking evaluation design for nutrition education programs. *Journal of Nutrition Education*, (13)3, 86-89.
- Kirk, Roger E. (1982). Experimental design: Procedures for the behavioral sciences. Belmont, CA: Wadsworth, Inc.
- Kirkpatrick, D. L. (1987). Evaluation of training. In R. L. Craig (Ed.), *Training and Development Handbook*, 3rd Edition. New York: McGraw-Hill.
- Kirkpatrick, D. L. (1983, November). Four steps to measuring training effectiveness. *Personnel Administrator*, 19-25.
- Koblinsky, Sally A. (1992). Evaluation of a nutrition education program for Head Start parents. Journal of Nutrition Education, 24(1), 4-12.
- Krein, T. J. and Weldon, K. C. (1994). Making a play for training evaluation. *Training and Development*, 48(4), 62-67.
- Laird, Dugan. (1985). Approaches to training and development. Reading, Mass.: Addison-Wesley Publishing Company, Inc.
- Lamme, L. L. and Hysmith, Cecilia. (1991, December). One school's adventure into portfolio assessment. *Language Arts*, (68), 629-640.
- Lederman, L. C. (1990, April). Assessing educational effectiveness: The focus group interview as a technique for data collection. *Communication Education (38)*, 118-127.
- Ludeman, Kate. (1991, November). Measuring skills and behavior. Training and Development Journal, 61-66.
- Madaus, G. F. and Stufflebeam, D. L. (1984). Educational evaluation and accountability A review of quality assurance efforts. *American Behavioral Scientist*, 27(5), 649-672.



849 Note S



- McKenzie, K. M. (1991, June). Evaluating instructors. Data Training, (30-32).
- Murphy, K. R. and Anhalt, R. L. (1992). Is halo error a property of the rater, ratees, or the specific behaviors observed? *Journal of Applied Psychology*, 77(4), 494-500.
- Murphy, Sheila E. (1987, June). Mirroring instructional technique: A process evaluation model. Training and Development Journal, 104-107.
- Newstrom, John W. (1987. July). Confronting anomalies in evaluation. *Training and Development Journal*, 56-60.
- Pulley, Mary L. (1994). Navigating the evaluation rapids. Training and Development, 48(9), 19-24.
- Roberts-Gray, Cynthia, Sparkman, A. F., Simmons, L. F., Buller, A. C., and Engquist, K. (1989). Evaluation of Texas' child-care-oriented nutrition education and training program. Journal of Nutrition Education and Training Program, 21(1), 16-24.
- Rosen, Ned. (1987, November). Employee attitude surveys: What managers should know. Training and Development Journal, 50-52.
- Rotondi, T. R. (1989, February). The anonymity factor in questionnaire surveys. *Personnel Journal*, 92-101.
- Schnitzer, Sandra. (1993, April). Designing an authentic assessment. *Educational Leadership*, 32-35.
- Smith, J. E. and Merchant, Sharon. (1990, August). Using competency exams for evaluating training. *Training and Development Journal*, 65-71.
- Society for Nutrition Education. (1984). An inventory of nutrition education evaluation instruments. Oakland, CA: Author.
- Stufflebeam, Daniel L. et al. (1980). An analysis of alternative approaches to evaluation. Educational Evaluation and Policy Analysis, 2(3), 5-20.
- Trapnell, Gail. (1988). A systematic approach to the evaluation of training and development programs. Atlanta: Trapnell, Wilkie & Associates.
- U.S. Department of Agriculture and U.S. Department of Health and Human Services (1990). Dietary Guidelines for Americans. Washington, D.C.: Author.

849 NoteS



- Van Velsor, Ellen and Wall, S. J. (1992, March). How to choose a feedback instrument. *Training*, 47-52.
- Wareing, Carol. (1990, February). Up close and personal: A model for supervision and evaluation. Supervision and Evaluation, (63), 247-250.
- Wentling, Tim L. (1980). Evaluating occupational education and training programs. Boston: Allyn and Bacon.
- Wilson, M. A., Harvey, R. J., and Macy, B. A. (1990). Repeating items to estimate the test—retest reliability of task inventory rating. *Journal of Applied Psychology*, 75(2), 158-163.
- Worthen, Blaine R. (1987). Evaluating educational and social programs. Norwell, MA: Kluwer Academic Publishers.
- Zemke, Ron and Gunkler, John. (1985, April). 28 techniques for transforming training into performance. *Training*, 48-63.







Appendices

- A. Guidelines For Developing Evaluation Instruments
- B. Data Collection Method: Case Study
- C. Data Collection Method: Participant Reaction Questionnaire
- D. Data Collection Method: Portfolio Assessment
- E. Sampling Protocol For NET Needs Assessment Surveys
- F. Excerpts from The Strategic Plan for Nutrition Education

NetS 628 NOO

Appendix A

Guidelines for Developing Evaluation Instruments

The development of instruments for evaluation involves the use of accepted design principles. Phillips (1991, pp. 82-87) identifies six important principles to be addressed during the development of evaluation instruments.

Validity

Validity is one of the most important characteristics of an evaluation instrument. A valid instrument measures what the evaluator wants to measure. The American Psychological Association has adopted four approaches in determining the validity of an instrument (Phillips, 1991, pp. 82-85).

- Content validity refers to how well the instrument represents the content of the program to be measured. This type of validity is insured when the important behaviors and information covered are represented in a balanced way in the instrument.
- Construct validity refers to the extent to which the construct to be measured is represented in the instrument. Constructs represent variables such as attitude, skill, or ability. Construct validity is a complex issue. One of the most efficient ways to determine construct validity is to have a group of experts review the instrument and determine whether it is an accurate measurement of the construct(s).
- Concurrent validity means the extent to which the results from administering an instrument agree with the results of other instruments used at approximately the same time to measure the same characteristics. For example, an attitude survey for CACFP Sponsors regarding nutrition education in Centers is administered to a sample of Sponsors. Another attitude survey is administered to the same group and both surveys obtain very similar results. The survey instrument would have concurrent validity.
- Predictive validity refers to the extent an instrument can predict future results or behaviors. This type of validity would be important for nutrition education programs where the evaluation question addresses predicting food related behavior (at school) using results of evaluation of a nutrition education program.

Phillips (1991, p. 85) suggests some simple guidelines for improving validity. These guidelines can and should be used when developing any kind of evaluation instrument.

- 1. Include an adequate number of appropriate items in the instrument. This is a question of balance since too few items compromise the validity and too many become cumbersome and could prevent some respondents from completing the form.
- 2. Make every effort to reduce response bias. People tend to say what they think their supervisors want to hear. Thus, it is important to design instruments in which people can give honest responses. Insuring confidentiality is important for many types of evaluation.
- 3. Administer the instrument objectively. Staff members or consultants who administer evaluation instruments should be trained to use the type of evaluation technique in a professional manner.





4. Recognize the weak link between attitude and behavior. This guideline is particularly important in nutrition education and food service training. It is important to recognize that a short training workshop of a few hours or a few days may positively influence attitudes, and even improve content knowledge, but is less likely to change behavior. Even though a person may indicate a certain attitude, his/her behavior may be inconsistent with the attitude. A nutrition education strand of a health curriculum taught for several hours over a six week period in one grade level may positively influence a student's attitude and interest in nutrition and may increase the student's knowledge of nutrition. However, such a program is unlikely to result in significant changes in food choices or in nutritional status.

Reliability

Reliability, another important characteristic of evaluation instruments, means that an instrument provides consistent measurements of an item. For example, if Summer Food Service Sponsors were administered a survey questionnaire regarding their attitude toward providing nutrition education materials to students at feeding sites and the same survey was administered several days later, a reliable instrument would obtain the same or very similar results.

There are four generally accepted procedures for ensuring reliability:

- II test/retest,
- ☐ alternate-form method.
- ☐ split-half procedure, and
- ☐ inter-item correlations.

An explanation of these procedures can be found in textbooks and journal articles listed in the References chapter of this Guide.

Ease of administration

An evaluation that is difficult to administer should not be used. Directions should be simple and easy to understand both for the people administering the instrument and for the respondents. A field test with a typical group of respondents can provide the evaluator with important feedback regarding the ease of administration.

Simplicity and brevity

The rule of thumb here is to be as brief as possible to cover the topic and cover it in the simplest way. Length does not equal effectiveness. Consider the level of readability of written instruments and the background of respondents. The readability level increases when sentences are long and multisyllabic words are used. To promote ease of readability when developing data collection instruments:

- □ Be brief.
- ☐ Keep it as simple as possible, and
- ☐ Leave as much white space as possible on each page.

A field test can provide valuable information regarding how well the evaluation designer has addressed this important consideration.



NoteS

Evaluation Guide for the Nutrition Education and Training Program

Economy

A good instrument is one that obtains the evaluation information specified by the objectives for the least cost. Costs are affected by design time, time for administration, time for analysis, costs of analysis, and time needed for developing the report. When data that are already available can be used for evaluation, of course, the cost is drastically reduced. The Data Review Method is described in the Appendix of the NET Needs Assessment Guide (1994) and offers a cost effective method of evaluation in certain situations.

As with needs assessment, the NET Coordinator must balance the cost of obtaining the evaluation information needed for program management decisions and accountability with the potential usefulness of the data obtained. The development of evaluation instruments can be time consuming and thus, costly. However, some types of evaluation are not costly at all. The overriding consideration is that an investment in evaluation is often the most critical investment in the program's future.

Note S

Appendix B

Data Collection Method: Case Study

Definition

The case study is an intensive, detailed analysis and description of a single program, class, institution, or phenomenon in the context of its environment.

Characteristics

A case study used as a data collection method for evaluation is:

<u>Specific</u> - focuses on events in one particular setting such as a class in management training, a specific group of new sponsors, particular types of instructors such as those who specialize in train-the-trainer workshops.

<u>Holistic</u> - tries to capture the totality of the phenomenon; looks at each part as well as the whole; looks at everything that makes up that entity.

<u>Longitudinal</u> - usually tells a story over time, although at times it may look at a "slice of life" or a given instance.

Qualitative - uses a variety of methods to collect information including interviews, observations, diaries, checklists, and data review or records analysis. The emphasis and focus is on the quality of the specific information, not the quantity. This method does not lend itself to "statistical numbers-crunching."

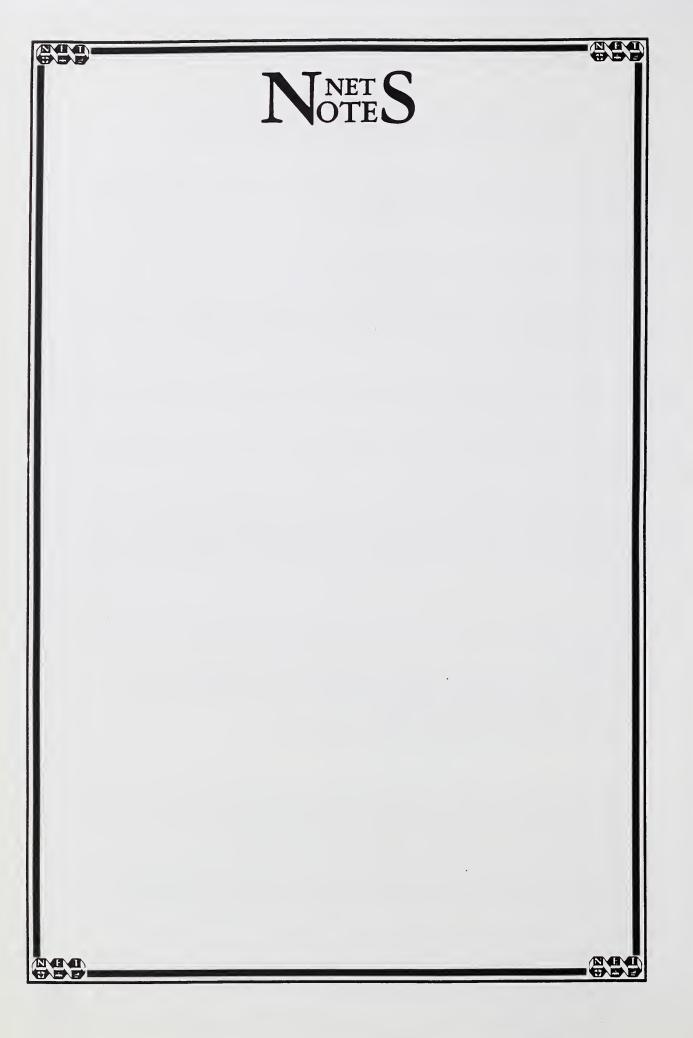
Advantages

- Problems can be formulated as the evaluation progresses since the evaluator is not bound to a priori questions.
- Data categories can be constantly modified making them more suitable to analysis.
- \square The distortion of variables that may be difficult to quantify can be reduced by the evaluator.
- ☐ Personal and immediate feedback is provided.
- Needs of diverse people and situations can be accommodated.

Limitations

- Although a case study data collection method may appear to be simple to conduct, it requires considerable skill in such areas as interviewing, scientifically sampling case study sites, and designing coding systems which allow for easy retrieval of information.
- The evaluator may fall into the trap of looking at everything and thereby losing focus. This creates a difficult situation in which to summarize pertinent information and findings.





- ☐ Case study reports tend to be lengthy since they include a chronicle of concrete activities or happenings.
- The evaluator cannot generalize his/her findings to other situations or groups.

Procedures

- 1. Set the boundaries by geographical areas, themes, and/or interests. Consider cost, time, and other constraints.
- 2. Determine the unit of analysis or the "thing" to be studied.
- 3. Select the site(s). Decide whether you will sample a group or evaluate the whole.
- 4. Establish initial contacts at the site. This may include seeking permission and gaining the confidence and confidentiality of the supervisors and others involved in the evaluation process.
- 5. Develop data collection procedures. These may include observation, interview, records analysis, and/or gathering data unobtrusively.
- 6. Develop a plan for organizing, coding, and storing data.
- 7. Train staff in the methodology to be used.
- 8. Implement the data collection procedures being careful to plan in advance, confirm interviews, and arrange for observations.
- 9. Analyze data and write the report.



Appendix C

Data Collection Method: Participant Reaction Questionnaire

Definition

The participant reaction questionnaire is designed to obtain the opinions of the participants regarding the various aspects of a presentation or training program. The questionnaire can be used during a training program (formative) or at the end of a training program (summative). It can be used to obtain feedback useful for input and process evaluation.

Characteristics

A well-designed participant reaction questionnaire provides the trainer or meeting planner with specific information for use in improving the program. Most questionnaires ask participants to provide feedback regarding their perceptions of the:

- Design of the training program,
- I Instructional techniques used by the instructor/trainer,
- Psychological climate set up by the instructor,
- ☐ Physical comforts, and
- Learning experience in totality (attitude toward the experience).

Procedures

- 1. Determine the information needed. Do not obtain information that is not needed, e.g., "nice to know." The information needed will be affected by whether the form is to obtain input evaluation information or process evaluation information. Oftentimes, a participant reaction questionnaire obtains data useful for both types of evaluation.
- 2. Decide the types of questions to be asked. Keep in mind that the type of question will have a major affect on the time needed to analyze the responses. Avoid openended questions with a group of more than 10 people. Use of ranking is also a difficult style of question to analyze.

Consider using:

- ☐ A checklist,
- ☐ Multiple-choice questions, or
- A rating scale.
- 3. Develop the questions. Write questions that are simple and straightforward. Avoid questions that can be confusing because of jargon, unfamiliar words, or phrasing. Limit each question to one idea.





Evaluation Guide for the Nutrition Education and Training Program

- 4. Test the questions. To the extent practical, obtain input from individuals who could be participants in the training where the questionnaire will be used. Revise the questions until they are clear.
- 5. Develop the completed questionnaire. Provide clear instructions both for the person administering the form and for the participants who will complete it.
- 6. Most participant reaction questionnaires are completed anonymously. If this is so, note that on the questionnaire.

An example of a participant reaction questionnaire is provided on the next page.

Participant Reaction Questionnaire

Workshop Title							
structorDate							
Please mark your response for items 1-25 in the approp	priate box. Do	not sign y	our name	·.			
Leave your completed questionnaire on the designate							
The Instructor							
	Strongly	Disagree	Agree	Strongly			
1. Demonstrated knowledge of the content.	Disagree			Agree			
2. Showed me how to apply the content to my job.	0	<u> </u>	0	Ö			
3. Was organized.	ō	ă	ă	ŏ			
4. Was easy to understand.	ō		ă	ō			
5. Demonstrated enthusiasm.	ā	ā	ā	ā			
6. Provided a variety of examples and demonstrations.	ā	ā	ā	ā			
7. Provided a variety of opportunities for participants' input.	ā		ā	ā			
8. Gave participants feedback during opportunities to practice.							
9. Encouraged participation during the training.							
10. Provided summaries of information periodically.							
11. Answered questions from the participants.							
12. Set up a comfortable climate for learning.							
13. Stayed on the time schedule.	_	_	<u> </u>	<u> </u>			
14. Used audio-visual and support materials appropriately.							
The Learning Experience							
	Strongly	Disagree	Agree	Strongly			
15 771	Disagree			Agree			
15. The amount of time was adequate for the material that was included.	u			3			
16. I can use the information in my work.							
17. The learning experience met my expectations.	<u> </u>		<u> </u>	- i			
18. I will use the workshop or course materials as		- i					
a reference.	_	_	_				
19. As a result of this experience, I will work to make some							
changes when I get back to my job.	_		_	_			
20. I would recommend this workshop/course.	0		0	0			
The Facility							
1100 1 4000000	Strongly	Disagree	Agree	Strongly			
21. Meeting place was convenient.	Disagree			Agree			
22. Meeting room was confortable.	<u> </u>	0	0.0	0 0			
23. Meals and/or snacks were enjoyable.	ö	ö	Ö	ň			
24. I could see and hear the instructor.	ŏ	ō	ō	0 0			
25. Materials helped me to learn.	ō	ō	ā	ā			
Please complete the following statements.							
26. The most positive aspect of this workshop/course was							
27. To improve this workshop/course the next time, I suggest							

Thank you for providing this important feedback.

Evaluation Guide for the Nutrition Education and Training Program





Appendix D

Data Collection Method: Portfolio Assessment

Definition

Portfolio assessment is one method of authentic assessment where learners assemble a portfolio or collection of documents which demonstrate their acquisition and application of the knowledge and skills defined by the performance objective(s).

Characteristics

A student portfolio may be an actual collection of documents from different sources or it may be the completion of several documents included in a workbook. It should include:

- ☐ Evidence of progress toward the performance objective(s),
- I Items that indicate transfer of learning from classroom to life or job,
- Things that convey learning activities or artifacts from lessons,
- ☐ Evidence of self-reflection and self-knowledge,
- I Items that tell the story of the journey, of the year, or the training program,
- I Things that show growth and change,
- I Items that demonstrate "my best work to date" and,
- Anything the student wants for whatever reason—after all, it is his/hers!

Procedure

This type of authentic assessment can be used effectively to document student performance (Valencia, et al, 1990). However, in the use of this type of assessment for process or product evaluation, "theory seems far in advance of practice" (Calfee and Perfumo, 1993). It involves the student assembling a collection of documents that, taken together, show the student's work. Historically, certain professionals, such as artists, fashion designers, architects, advertising agents, and others have assembled a portfolio to illustrate their body of work. This same concept is applied in portfolio assessment. Portfolio assessment is authentic in that it shows an application that focuses on how an individual can use the knowledge and skills acquired.

1. Have the student compile several documents into a portfolio that demonstrates acquisition of knowledge and skills congruent to the performance objectives. Such portfolios serve to document the authentic assessment of student performance and help to integrate assessment with instruction. The portfolio of student work can provide a basis for ongoing dialogue between the teacher and the student, the student and the parent, the parent and the teacher, and subsequently between the student and a teacher in another school, as well as, perhaps, between the student and an employer. Many schools across the country are, in fact, moving to the use of student portfolios in place of student transcripts.



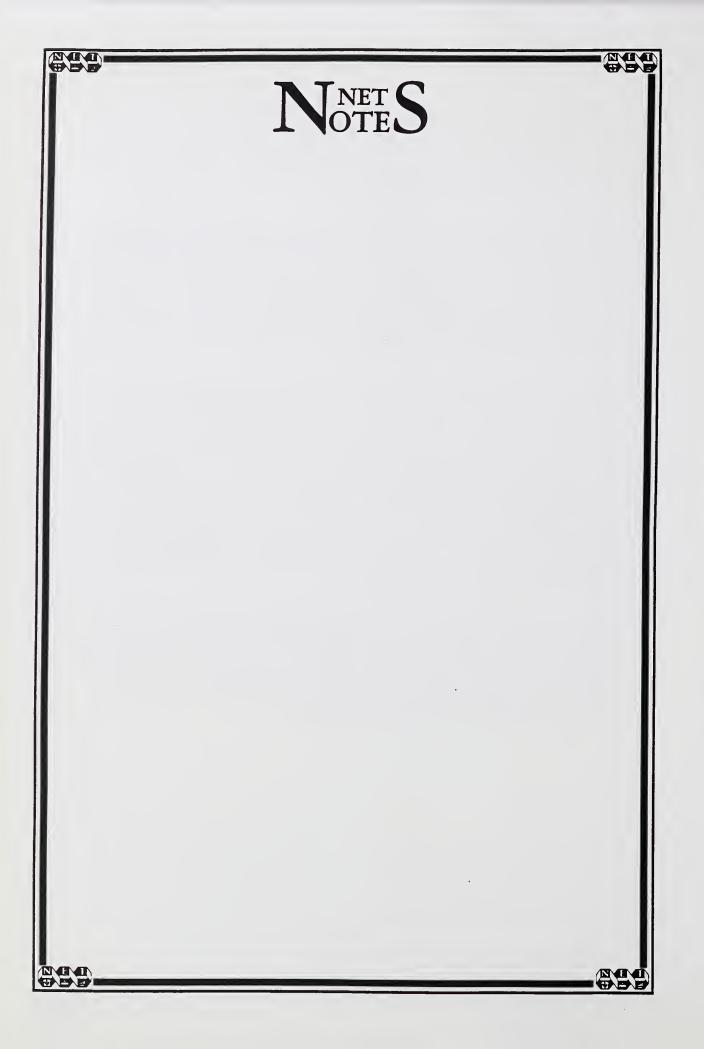
This evaluation technique can be used effectively for adult training situations and is being used more frequently. Realistically, in order for portfolio assessment to be used successfully in adult training experiences, the training must occur over a period of time that allows for the learner to assemble several evidences of acquisition and application of the training content. The design considerations of portfolio assessment are the same for K-12 students and adult learners.

2. Assess the assembled portfolios using pre-determined criteria based on performance objectives (measurable outcomes). The individual responsible for the assessment of the student-assembled portfolio must be given a proper orientation and training experience. The assessment of the portfolios in K-12 classrooms may involve the teacher, the student, classmates, parents, community members, or known experts. The assessment of the portfolios in adult training may include the instructor, the adult learner, peers, a supervisor, or known content experts.

One of the concerns about using this type of assessment as a method of process or product evaluation is the issue of validity and reliability. Validity can be assumed because this assessment process is based on student assembled portfolios demonstrating authentic application of new knowledge and skills. However, reliability is a different matter. In one study reported with language teachers, a panel was used to establish reliability (Calfee and Perfumo, 1993). Two or more teacher-judges scored each student portfolio to establish consistency with the stated performance objectives. Calfee and Perfumo report that this practice is rare. When portfolio assessment is used in the classroom or a training situation, the instructional designer should pay careful attention to the issue of reliability.

Schools report using portfolio assessment as one of several measures of student achievement. Such an approach is also appropriate for use in adult training situations. It was used in the design of a five-day orientation program for new school food service managers in Mississippi titled, *DECIDE to Succeed: An Orientation for New Managers* (1995, Office of Child Nutrition, Mississippi Department of Education) and is also being used to assess teachers in a national program.





Appendix E

Sampling Protocol for NET Needs Assessment Surveys

Prepared for the 14-State NET Needs Assessment Consortium
by
Jack Hamilton, Ph.D.
Lumina Training Associates, Birmingham, Alabama
November 1994

Introduction

A sample is a subset of a population observed for purposes of making inferences about that population. The quality of a sample is based on the degree to which it is representative of the population. A sample is considered to be representative if the characteristics of the sample are generally the same as those of the population from which it was selected.

Although there are a variety of sampling procedures available to researchers, the discussion which follows is limited to *probability sampling*. The basic principle of probability sampling is that a sample will be representative of the population from which it is selected if all members of the population have an equal chance of being selected in the sample. It is helpful to think of probability sampling as involving four separate steps:

- 1) identification of the population and selection of an appropriate sampling frame,
- 2) determination of the sampling procedures to be used,
- 3) determination of the minimum sample size acceptable for the research purpose, and
- 4) selection of the sample.

It should be noted however, that there is considerable variation available within each step depending on the number and distribution of the populations to be surveyed.

Populations and Sampling Frames

The researcher must begin with a careful specification of the population to be sampled, such as the population of school food service district directors. Once this population has been identified, the researcher must obtain or create a sampling frame. The *sampling frame* is an actual list of the elements or individuals in the population from which the sample will be selected.

The characteristics of the sampling frame are largely determined by the form and extensiveness of available records. The simplest sampling frame is an available list of all members of the population. The State Department of Education Directory usually provides a list of the school food service district directors and can be used as a sampling frame for this population. This same directory also provides a list of school districts and schools, and thus can be used as the sampling frame for other NET-targeted populations. The State Agency which administers the Child Nutrition Programs has lists of other populations including the Child and Adult Care Food Program popu-





lations. An available list of all members of a population can be used to select a sample in a single stage. However, when such a list is unavailable or non-existent the researcher must create such a list from other available information, often involving several different stages in the selection of the sample. It should be noted that probability sampling is not possible if a sampling frame does not exist or cannot be created.

Sampling Procedures

Two requirements must be met in order to select a probability sample: (1) the availability of a complete list of the elements or persons in the population, and (2) the consistent application of randomization procedures.

A simple random sample can be utilized when the sampling frame is comprised of a complete list of the population to be sampled. When such a list is available, the researcher can draw a simple random sample in one of two ways. First, if the sampling frame is on a computer, then a computer program can be written to generate a random sample directly from the list. Secondly, if the sampling frame is only available on a hard copy list, the researcher can assign numbers to each element of the population, and select the sample by referring to a "Table of Random Numbers." (A table of random numbers can be generated by computer, or the researcher can find such tables in the appendices to most textbooks in statistics and social research.)

The selection of a simple random sample requires the researcher to: (1) assign a numerical identification to each element of the population, (2) specify the sample size, and (3) use random numbers which correspond to the assigned numbers to select the sample. Given a list of 100 CACFP sponsors, for example, a sample of 20 sponsors can be easily selected.

First, the researcher assigns each sponsor a sequential number from one to 100. Secondly, the researcher obtains twenty random numbers between one and 100 by: (1) generating twenty different random numbers, each between one and 100, by computer, or (2) selecting the first twenty numbers between one and 100 from a table of random numbers. Thirdly, the researcher selects the sample of twenty by matching each random number with the number assigned to the population. Thus, if the first random number is 55, the 55th person on the list of sponsors is selected.

Although the use of a table of random numbers is quite simple, the actual process of selecting a simple random sample can prove to be tedious and laborious, especially if the population listed on the sampling frame is quite large. Consequently, *systematic random sampling* is used and preferred by most researchers.

Systematic random sampling recognizes that a systematic sample will be equivalent to a simple random sample, if the sampling frame used is essentially random. To select a systematic random sample the researcher must:

- 1) specify the sample size,
- 2) determine a sampling ratio (sample size divided by the population size), and
- 3) randomly select the first element or person.



For example, a researcher who wishes to survey 10 people from a population of 100 people will have a 1/10th sampling ratio. Consequently, all the researcher must do is randomly select the first case (perhaps using a table of random numbers), and thereafter select every tenth name of the list.

In many studies the researcher wants to assure that different kinds of conditions or groups are also represented in the sample. In such situations the sampling procedure must be modified to assure that the sample includes representatives of such groups ("strata"). Thus, whether the researcher is drawing a simple random sample or a systematic random sample, the sample procedure may include the selection of a stratified random sample.

In a stratified random sample the researcher divides the individuals or groups into categories, and then independent samples are selected from each group or strata. Proportional stratified sampling, in which the sample from each group is proportional to the size of the group in the population, is generally preferred in order to assure a more representative sample than would be obtained by use of only a random sample or a systematic random sample. Proportional stratified sampling is appropriate for most NET needs assessment surveys. The student enrollment figures for school districts can be used to determine the strata for samples of school food service personnel, principals, curriculum directors, or other groups in schools.

Determining Sample Size

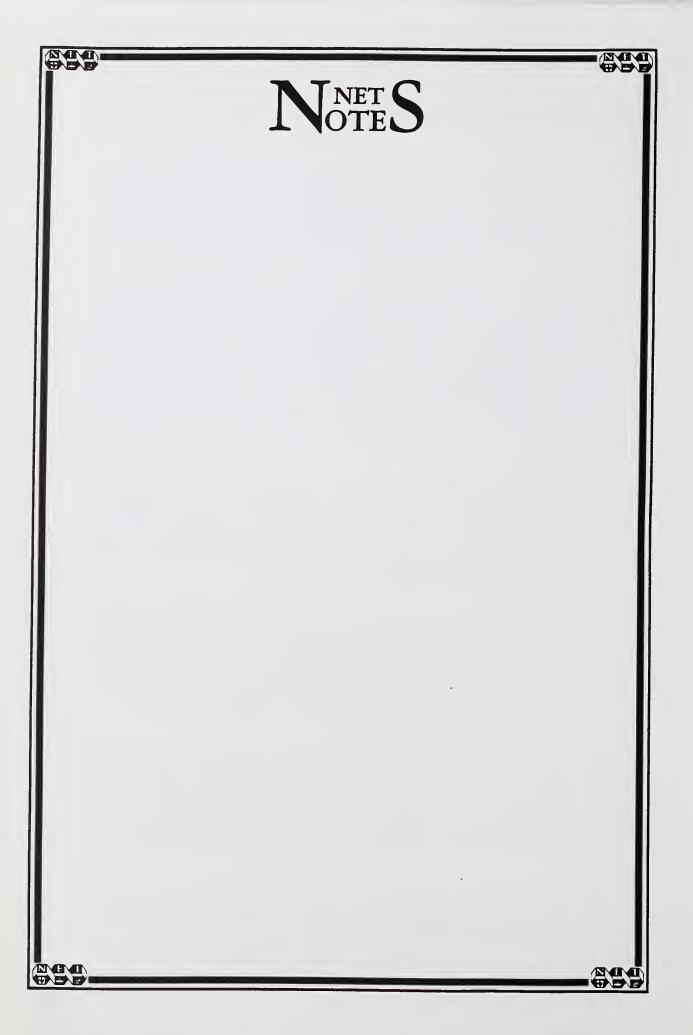
The researcher should be aware that the size of the population has relatively no relationship to the size of the probability sample to be selected. Instead, the size of the sample is largely a function of the desired level of accuracy. Although a discussion of probability theory is beyond the scope of this overview, an understanding of a few basic assumptions are of value to an understanding of statistical analysis and the determination of sample size.

Many of the statistical procedures for determination of sample size as well as data analysis are derived from the concept of a normal distribution. The normal distribution is commonly depicted as the traditional and perfect "bell-shaped curve." However, this distribution is a theoretical distribution which is defined only by mathematical rules. In this theoretical normal distribution the horizontal axis is continuous and represents all possible values between plus (+) and minus (-) infinity. The vertical axis represents the probability density for each possible value.

The conception of a normal distribution permitted statisticians to mathematically construct standardized tables of probabilities. Further, the existence of a standardized normal distribution permitted statisticians to derive mathematical equations which provided for a variety of ways to convert sample data to standardized measures which, when compared with the standardized normal distribution permits the researcher to identify the probability of obtaining such results from chance expectation.

The normal distribution is an excellent approximation of the sampling distribution of the population mean. In addition, the sampling distribution of the sample mean approaches a normal distribution as the sample size increases. Thus, with a sufficiently large sample, the normal distribution provides an excellent approximation to the exact sampling distribution.





A major advantage of probability sampling is that probability theory permits the researcher to estimate the representativeness of the sample. Any sample estimate may include some margin of error. This possible error is typically referred to as the *sampling error*, and refers to the *statistical error* which may arise from any effort to represent a population with a sample. A description of the sampling error for a given sample contains two measures, the confidence level and the confidence interval.

The *confidence interval* refers to the degree of sampling error that the researcher is willing to accept. A sampling error of 5%, for example, means that the error may be as much as 5% greater than, or 5% less than, the true population mean. This sampling error can be expressed as a confidence interval ranging from +5% to -5%.

The *confidence level* is used in conjunction with a specified confidence interval, and serves as a measure of the degree of confidence in the likelihood that the sample is within the specified error range. Statistically, the 95% confidence level (.95) means that if a researcher selected 100 different samples, 95 of the samples would fall within the specified confidence level.

Table 1., "Minimum Sample Size and Sampling Error, by Confidence Level and Confidence Interval" which follows, is provided to illustrate how sampling error is related to sample size. As Table 1. illustrates, a researcher who is willing to accept the conventional 95% confidence level that the sampling error is within a 4% sampling error would need to select a sample of 600 people. In order to be 99% confident that the sample is within 4% of the population mean, the researcher would need to increase the sample size to 1,037. The specified sample size, in either case, would be appropriate whether the population being sampled was a city, a state, or the United States of America. (The Gallup Poll, for example, typically interviews about 1,200 people as an estimate of a population in excess of 200-million people. Table 1. reveals that Gallup can be 95% confident that the sample mean is within a 3% sampling error of the true population mean.)

Table 1.

Minimum Sample Size and Sampling Error
by Confidence Level and Confidence Interval

Acceptable Confidence Level		
.95	.99	
9,604	16,587	
2,401	4,147	
1,067	1,843	
600	1,037	
384	663	
267	461	
196	339	
	.95 9,604 2,401 1,067 600 384 267	

SOURCE: Adapted from Charles Backstrom and Gerald Hursh. (1963). Survey Research. Minneapolis: Northwestern University Press, pp. 28-33.

It is important to remember that the size of the sample is largely a function of the desired level of accuracy, and is not greatly influenced by the size of the population. The sample sizes reported in Table 1. reveal that, regardless of the size of the population, extremely large samples are rarely needed. In fact, although generally considered too small for such precise and rigorous statistical analysis as required in the formal testing of a specified hypothesis, the researcher who draws a probability sample of 100 people can be 95% confident that the results are within a 10% sampling error. Since NET needs assessment data are most likely to be used to make general management decisions, a NET researcher may tolerate a larger sampling error and be comfortable with a smaller sample size. It is recommended for NET needs assessment surveys that, whenever the population group includes 200 people or less, the entire population be included in the survey. The use of the entire population eliminates all sampling error and, of course, makes it unnecessary to select a sample from that group.

Selecting the Sample

The overview of the first three steps in the sampling process reveals that the process of selecting a sample can be undertaken in a variety of ways. However, once the researcher identifies the sampling frame, defines the sampling procedure, and specifies the size of the sample, then the actual selection of the sample is quite simple. Beginning with a list of the population (the sampling frame), the researcher merely follows those procedures in order to prepare a list of the sample population.





The sampling procedures described above suggest that the preferred sampling procedure for the NET needs assessment surveys consist of the selection of a systematic stratified random sample. However, in order to simplify the description of the sampling procedures, the overview of the four steps was presented in terms of a single sample. Unfortunately, the selection of the sample does not always proceed as smoothly as the four steps indicate. Consequently, there are several useful considerations that may help the researcher deal with certain problems as they arise.

First, it is helpful to begin a survey design with an awareness that the researcher will often be drawing a series of samples rather than a single sample. This may occur as a result of a number of factors, including the possibility that:

- 1) different groups or strata may be found within the population,
- 2) each such group may have a different sampling frame in a different format,
- 3) the questionnaires to be administered may not be identical for each organizational unit, and
- 4) in many cases the data collected for each group within the sample will be coded and recorded in a separate file for purposes of data analysis.

Similar agencies in different States, as well as different agencies in the same State, do not always collect the same information on the same form. Thus, for example, in attempting to select a sample of "School Food Service (SFS) Managers," the researcher in "State A" may be fortunate enough to find that the State Agency has a list of the names and addresses of the SFS Managers in every school district. A random sample of SFS Managers can be drawn directly from such a list.

In contrast, a researcher in "State B" may find that there is no list of all the SFS Managers in the State. Instead, the available data are in the form of a list of all public and private schools and organizations that prepare food under the Child Nutrition Program. This is a more complicated situation, and will require the researcher to indirectly select a sample of SFS Managers by selecting a sample of school districts and schools.

Secondly, it should be noted that differences in the available sampling frames not only influence the way the sample is selected, but will also affect procedures for administering the survey. In the two situations described above, the researcher in "State A" can mail the survey questionnaires directly to the SFS Managers, but the researcher in "State B" must mail the questionnaire to the school district Food Service Directors, relying upon them to distribute the questionnaire to the Managers.

Thirdly, the researcher must realize that the characteristics of the available sampling frame or list will also influence the selection of the sampling procedures to be used. More importantly, the researcher should be aware that a different sampling procedure may be used with different populations. Thus, were one stratum or group has an extremely large population, a systematic random sample might be selected. This procedure, for example would be suited to the selection of a sample of Family Day Care Home Providers. In contrast, where another group has an extremely small population, a simple random sample might be preferred. This procedure would be more appropriate for the selection of a sample of Child Care Center Sponsors.

Fourthly, although a stratified sample is preferred, the researcher must be aware that the same



strata may be defined differently for various groups within the population. The researcher may be interested in differences between public schools and private organizations that provide meals. The researcher may also be interested in differences with respect to the number of meals served, distinguishing between large, medium, and small programs based on the number of meals served. In public school Child Nutrition Programs, "large" may be defined as over 1,000 meals, whereas in child care centers, "large" may be defined as over 50 meals.

The research problem in this situation is the recognition that no one definition of the strata is appropriate for all populations being surveyed. The best advice to the researcher in such a situation is to apply to each population: (1) the same definition or categories for each strata examined (such as small, medium, and large), and (2) equal intervals of measurement for each category, based on the largest and smallest numbers for the respective population.

Nevertheless, despite such difficulties, the use of stratified sample serves to improve the representativeness of the sample because it reduces the degree of sampling error. There are two potential sources of variation in a random sample. Sampling error may be found within each strata as well as between strata. The stratified sample serves to effectively eliminate the sampling error between strata.

Finally, when confronted by a problem situation that requires a decision about the sampling process, the researcher should turn to the experts in survey research and statistical analysis. Numerous texts are readily available to assist the researcher, and a few examples are provided in the Selected Bibliography for this paper. Any problems in sample selection that cannot be readily resolved from the available and extensive literature, should be resolved by the principle of random selection. Such questions are best solved by "flipping a coin" because only a random decision is free of the possible bias of the researcher.

Selected Bibliography

Numerous books and articles are available to assist in all stages of research. The beginning researcher may consult an introductory textbook in research, such as Earl Babbie, Survey Research Methods (Wadsworth, 2nd ed., 1990), or The Practice of Social Research (Wadsworth, 6th ed., 1992), both of which include a Table of Random Numbers in the Appendix. Researchers interested specifically in the procedures for selecting a sample might refer to the classic works of Leslie Kish, Survey Sampling (New York: John Wiley, 1965); Hubert Blalock, Social Statistics (Minneapolis: Northwestern University Press, 1963). The advanced researcher who wishes to understand the intricacies of statistical analysis and probability theory, can trace the mathematical derivations of most statistical equations in William Hays, Statistics (New York: Holt, Rinehart, and Winston, 1963).

Copyright © 1994 by C-NET Needs Assessment Consortium (Used with permission.)



Appendix F

Excerpts from The Strategic Plan for Nutrition Education

United States Department of Agriculture Food and Nutrition Service Nutrition and Technical Services Division 1993

Edited by: R. Jane Mandell

The Strategic Direction

In the year 2000, nutrition education will be a major educational component of all child nutrition programs and offered in all schools, child care facilities, and summer sites as part of health education, as part of an interdisciplinary curriculum, or through other means. Nutrition education will be supported and promoted by State and local personnel working in partnership with the children and their families as part of the total school environment or child care setting. Depending on available resources, State agencies will conduct formal needs assessments and evaluations as the basis for program development, improvement, and accountability. Nutrition education staff at State offices, USDA Headquarters, and USDA Regional offices will have specific roles and responsibilities in the administration of the nutrition education program. NET will have extensive alliances and partnerships with public and private concerns at the local, State, and national levels. Through these relationships, NET will obtain funding and resources, identify research agencies in nutrition education and food management, identify technical assistance resources, develop evaluation strategies, and work cooperatively with all groups that promote nutrition education and service of nutritious meals and snacks for our nation's children.

The Goals

Nutrition Education and Training

Goal 1

Increase to at least 75% the proportion of the Nation's schools that provide nutrition education as part of health education, as part of an interdisciplinary nutrition education curriculum, or through other means to children and parents by the year 2000 as documented by data submitted to FNS or by other sampling methods.

Goal 2

Increase to a least 50% the proportion of child care and summer program facilities/sites that provide nutrition information/education to children, caregivers or parents by the year 2000 as documented by data submitted to FNS or by other sampling methods.



Goal 3

Increase to a least 50% the proportion of food service personnel/providers in Child Nutrition Programs who have received nutrition information and/or training (including food service training) by the year 2000 as documented by data submitted to FNS or by other sampling methods.

Nutritious Meal Service

Goal 4

Increase to at least 90% the proportion of school lunch and breakfast services with menus that are consistent with the nutrition principles in the Dietary Guidelines for Americans by the year 2000 as documented by data submitted to FNS or by other sampling methods.

Goal 5

Increase to 90% the proportion of child care and summer food services with menus that are consistent with the nutrition principles in the Dietary Guidelines for Americans by the year 2000 as documented by data submitted to FNS or by other sampling methods.

Leadership

Goal 6

Increase to at least 75% the proportion of States that incorporate applicable goals and strategies from the NET Strategic Plan into their NET State Plans by the year 2000 as documented by data submitted to FNS.

Goal 7

Increase to at least 75% the proportion of States that use national guidelines to conduct needs assessments by the year 2000 as documented by data submitted to FNS.

Goal 8

Increase to at least 75% the proportion of States that use national guidelines to evaluate NET projects by the year 2000 as documented by data submitted to FNS.

Goal 9

Increase to at least 75% the proportion of nutrition education personnel at the State, regional and national levels with roles and responsibilities that are consistent with national guidelines by the year 2000 as documented by data submitted to FNS.

Goal 10

Establish partnerships with at least 10 allied nutrition, health, food management, or education organizations; public or private agencies or groups; or others interested in child nutrition issues by the year 2000.







Glossary

Attitude A feeling or emotion toward a fact or state.

Audience The group of participants who are targeted for a program or project.

Context The situation, the environment, the milieu.

Context evaluation Evaluation of the situation or environment in which a program or project operates.

Criteria Standards for making a qualitative or quantitative judgment.

Data collection method A technique used to obtain data. The data may be used for needs assessment or for evaluation.

Descriptive statistics Statistics that help describe certain characteristics of a research sample. Two basic types of measures used are measures of central tendency and variability (dispersion).

Evaluation plan The systematic, written procedure describing how each State Plan objective, or program/project objective will be evaluated. The plan should include evaluation questions, the data to be collected, the standard to assess acceptability, the evaluation method/procedure to be used, a work schedule, and an identification of responsibility.

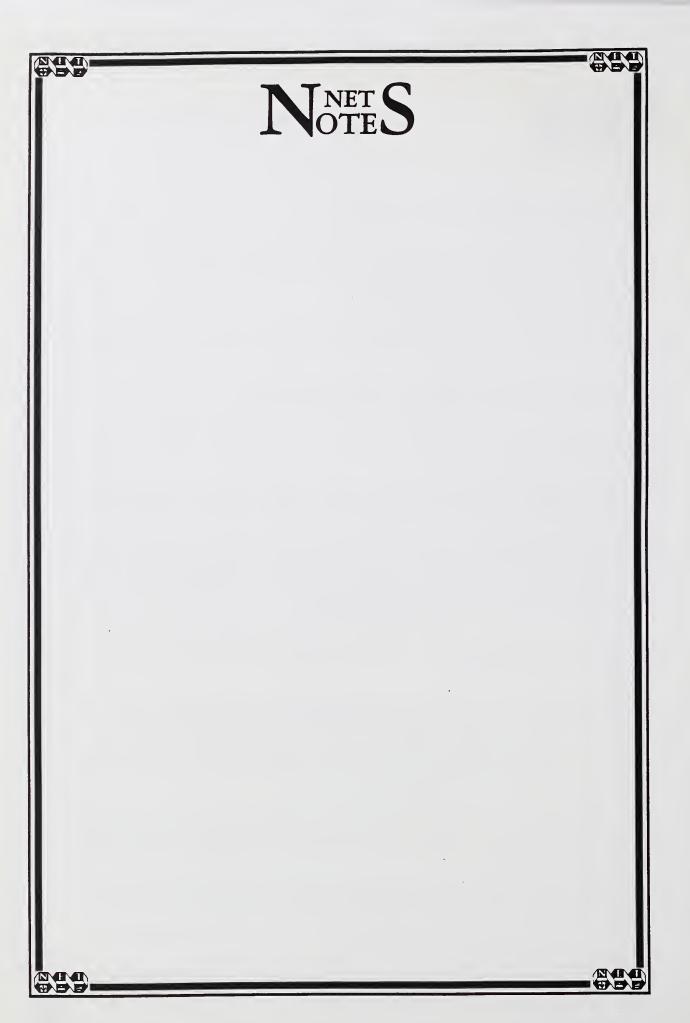
Evaluation type One of the five types of evaluation comprising the CIPPI model: context, input, process, product, and impact.

Formative evaluation Evaluation which occurs during the planning and implementation of a program or project. Context, input, and process evaluation may be used formatively. In fact, using these three types of evaluation as formative evaluation provides the program planner data on which to base decisions to make corrections in a program or project during the planning and implementation phase, rather than waiting until it is completed.

Frequency distribution Grouping a large mass of data into different classes and then determining the number of observations that fall in each of the classes. An arrangement of data in tabular form is called a frequency distribution.

Impact evaluation Evaluation of the cumulative differences a program or project makes; "the bottom line."







Inferential statistics Comprises those methods concerned with the analysis of a subset of data leading to predictions or inferences about the entire set of data.

Input evaluation Evaluation which assesses the appropriateness and adequacy of all resources that go into the design and implementation of a program or project.

Mean A measure of central tendency which is determined by the simple arithmetic average.

Median A measure of central tendency that is equivalent to the 50th percentile.

Mode A measure of central tendency that is the most frequently observed measure.

Opinion A generally held view; a belief stronger than an impression and less strong than positive knowledge.

Perception A mental image; discernment.

Population The total group or total number of observations with which a researcher is concerned.

Process A course of action or a series of steps leading toward an end.

Process evaluation Evaluation which provides information about program implementation and whether learners are meeting interim performance objectives.

Product evaluation Evaluation which focuses on the attainment of performance objectives after the completion of the program.

Prospective Relating to or effective in the future.

Proximity Soon forthcoming, very near.

Retrospective Relating to or affecting things past.

Qualitative data Data which are categorical.

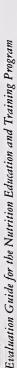
Quantitative data Data which are numerical.

Sample A subset of a population.

Summative evaluation Evaluation used at the end of a program or project. Context, input, process, product, and impact evaluation may be used as summative evaluation after a program or project has been completed.



829 NetS





A-P-B-I-E5,6 Authentic performance Case Study63, 64 CIPPI Model23, 24, 25, 26, 27 Closed Loop Assessment-Planning-Budgeting-Implementation-Evaluation Cognitive Response Method......35 Common statistical tests20 Context evaluation......23, 26, 33, 34, 35 Data collection......14, 15, 16, 17, 19 Ethnography......33, 34, 35, 37 Formal evaluation9 Gantt chart17, 18 Impact evaluation25, 26, 50, 51, 52 Inferential treatment of data19, 20

Index

Input evaluation23, 37, 38, 39
Interim performance
Learning taxonomy (a learning continuum)
Monitoring
Needs assessment5, 6, 37, 38, 62
Participant reaction questionnaire
PERT chart
Planning
Portfolio assessment44, 49, 68, 69
Process evaluation24, 40, 41, 42, 43, 44,
Product evaluation24, 45, 46, 48, 49
Proximity of outcomes10, 11, 49, 51
Qualitative evaluation methods9, 10, 35, 36, 37, 40, 41, 51
Quantitative evaluation methods9, 10, 37, 40, 41, 51
Reliability61
Sampling frame70

Informal evaluation9

Sampling procedure70
Sample size70, 73, 74
Simple random sample71
Strategic Plan for Nutrition Education
Stratified Random Sample72
Systematic random sample71
Systematic stratified random sample75
Summative evaluation7, 8, 25, 45, 65
Validity60, 61